Cadre discret stabilité au bruit et influences.

Chapitre 4

Introduction générale dans le cadre discret

L'analyse booléenne a connu un grand essor durant les 30 dernières années.

Les fonctions booléennes (i.e. à valeurs dans $\{0,1\}$) sont peut-être les objets les plus basiques en informatique théorique. Au-delà de cette branche, elles interviennent de manière cruciale dans de nombreux autres domaines des mathématiques, en combinatoire, théorie des graphes aléatoires, physique statistique mais aussi en théorie du choix social, théorie des espaces de Banach et géométrie gaussienne via le théorème de la limite centrale. Nous présentons, de manière non exhaustive, quelques domaines dont les problématiques ont des liens avec le cadre de notre travail.

Parmi les domaines étudiés, on peut citer en premier lieu la "théorie du choix social", une branche issue de l'économie, via l'étude de l'organisation d'un scrutin. L'issue d'une élection à deux candidats, disons 0 et 1 peut être modélisée par une fonction booléenne, dite fonction de choix, qui détermine le vainqueur. Au suffrage universel, la fonction de choix correspond à ce qu'on appelle la fonction Majorité. Lorsque l'issue ne dépend que de la volonté d'une seule personne, la fonction correspondante est celle qui associe une coordonnée. Ces fonctions sont appelées en conséquence fonctions dictateurs. D'autres systèmes de votes, qui respectent un principe moral de neutralité - ceux où aucun votant n'a trop d'influence sur le résultat - sont possibles. Citons l'exemple des majorités récursives qui régit l'élection présidentielle aux États-Unis, ou encore celle de majorités à poids.

Une élection peut également faire intervenir plus de deux candidats. Cette dernière situation a été étudiée dès le XVIIIe siècle par le marquis de Condorcet. Ce dernier énonce un célèbre paradoxe : il est toujours possible au suffrage universel d'obtenir un "issue irrationnelle", c'est à dire une issue où le premier candidat est préféré au deuxième, qui est préféré au troisième lui-même étant préféré au premier. An milieu du XXe siècle, l'économiste Arrow a démontré que seuls le choix d'une seule personne évite la possibilité d'une issue irrationnelle - c'est le théorème d'impossibilité d'Arrow [Arr]. Au début des années 2000, Kalai [Kal] donne une preuve théorique de ce théorème se basant uniquement sur l'analyse de Fourier discrète. Il prouve au passage que la probabilité d'obtenir une issue irrationnelle est minimisée pour le suffrage universel parmi les "bons" systèmes de votes, c'est-à-dire ceux tels que personne n'ait une grande influence sur le résultat.

En second lieu, on peut citer le domaine des graphes aléatoires. Sur de tels graphes, certaines propriétés sont modélisées par des fonctions booléennes. Par exemple, pour prendre un modèle simple, considérons un graphe à n sommets sur lequel chaque arête a une probabilité $p \in (0,1)$ d'être présente ou non. Alors plusieurs propriétés sur les arêtes peuvent être modélisées par des fonctions du cube discret $\{0,1\}^{\frac{n(n-1)}{2}}$ muni de la mesure biaisée ν_p. Par exemple, y-a-t-il une composante connexe dans le graphe ? Est-ce que le graphe contient un k-cycle ? Ces propriétés citées sont monotones par rapport à p. Un fait récurrent est de considérer une suite graphes dont le nombre de sommets tend vers l'infini. Est associée à cette suite de graphes une suite de fonctions booléennes $(f_n)_{n\geq 0}$ définies $\{0,1\}^{\alpha_n}$ pour α_n tendant vers l'infini. On étudie le comportement asymptotique de certaines de ces fonctions. Au milieu des années 90, Friedgut et Kalai [F-K] ont étudié les phénomènes de seuil de ces propriétés de graphes monotones pour n. Leurs travaux s'appuient sur des résultats fondateurs sur les fonctions booléennes du cube que nous présenterons.

Enfin, dans un autre domaine, en théorie de la percolation, de nombreuses propriétés sont modélisées par des fonctions booléennes. Sur le réseau $\mathbb{Z}^2$, considérons par exemple le rectangle $[n]\times[n]$. Mettons des arêtes sur le réseau avec probabilité $1/2$. À l'événement qui correspond à une percolation, c'est-à-dire un chemin de gauche à droite (ou de bas en haut) est associé naturellement une fonction booléenne de $\{0,1\}^{n^2} \to \{0,1\}$. L'étude théorique des fonctions booléennes et son application en percolation a été mis en lumière par Benjamini, Kalai et Schramm [B-K-S]. Ces auteurs mettent en lien deux notions centrales de notre thèse, celle d'influence d'une fonction booléenne et celle de sensibilité au bruit.

Une référence introductive, complète et récente sur l'analyse des fonctions booléennes est l'ouvrage de Ryan O'Donnell "Analysis of Boolean functions" [OD]. En ce qui concerne l'utilisation des fonctions booléennes en percolation, nous renvoyons le lecteur aux notes de Garban et Steif [G-S]. Pour ce qui est de l'utilisation des influences, nous renvoyons le lecteur au survey de Kalai et Safra [K-S].

La grande majorité des travaux effectués sur l'analyse des fonctions booléennes s'appuie sur l'analyse de Fourier sur le cube discret, devenue un outil indispensable. L'analyse de Fourier sur le cube commence au début du XXe siècle avec les travaux de Walsh et Paley qui construisent une base orthonormée sur le cube discret, désormais appelée base de Walsh ou Fourier–Walsh. Le résultat fondateur et sous-jacent à tous les développements récents date cependant du début des années 70. Ce sont les travaux de Bonami [Bon] (publiés en francais et semble-t-il passé d'abord inaperçu sur le plan international) puis Beckner [Bec] qui aboutissent au théorème d'hypercontractivité de qui est depuis appelé semi-groupe de Bonami–Beckner, un élément fondamental de l'analyse Booléenne. Notons que ce résultat est postérieur au premier théorème d'hypercontrativité du semi-groupe d'Ornstein Uhlenbeck de Nelson [Nel] (cependant incomplet) mais qu'il est plus fort au sens où, par théorème de la limite centrale, il le contient - comme remarqué par Gross.

Il faut cependant attendre la fin des années 80 pour voir ce théorème prouver sa force dans l'étude théorique des fonctions booléennes, avec les travaux de Kahn, Kalai et Linial [K-K-L] puis ceux de Talagrand [Tal1]. Ces travaux ont joué un rôle prépondérant dans l'analyse des phénomènes de seuil étroits dans les graphes aléatoires. Ils jouent également un rôle prépondérant dans un grand nombre de domaines en informatique théorique. Nous convions le lecteur à la discussion introductive de l'article [O-W] pour des références plus précises à ce sujet. L'hypercontractivité, notion rencontrée dans le cadre continu, joue un rôle récurrent dans le domaine de l'analyse des fonctions booléennes et est un élément à la base d'une très grande partie des résultats que nous citerons, y compris les nôtres.

Le but de la section suivante est d'introduire les concepts basiques d'analyse de Fourier discrète sur le cube. Ceci permet alors d'établir le théorème d'hypercontractivité dû à Bonami puis Beckner. Nous discutons ensuite de la généralisation à des espaces discrets plus généraux.

Dans les sections qui suivent, nous définissons les deux notions principales de notre travail, celles d'influences et de stabilité au bruit. Nous faisons un court rappel historique des résultats établis en lien avec ces notions. Nous nous concentrons ensuite sur l'analogue discret du théorème de Borell concernant la stabilité au bruit. Enfin, nous présentons nos travaux sur le critère quantitatif entre stabilité au bruit et influences.

4.1 Analyse de Fourier discrète sur le cube

Cette section présente les élément basiques de l'analyse de Fourier discrète sur le cube. Le cube discret n dimensionnel est l'espace à deux points tensorisé n fois. Par commodité, il est parfois préférable de le définir comme $\{-1,1\}^n$ ou comme $\{0,1\}^n$. Par ce qui concerne l'analyse de Fourier, il est commun d'utiliser $\{-1,1\}^n$ comme définition.

4.1.1 Base de Walsh

Soit $(\{-1,1\}^n, \nu)$ le cube discret muni de la mesure uniforme. Alors $L^2(\{-1,1\}^n, \nu)$ est un espace euclidien où le produit scalaire est le produit scalaire usuel $\langle\cdot,\cdot\rangle_{L^2(\nu)}$.

À la suite des travaux de Walsh, une base orthonormée de cette espace est connue. Il s'agit de $(W_S)_{S\subset[n]}$ où $W_S(x) = \prod_{i\in S} x_i$. En effet, par indépendance des x_i, $1 \le i \le n$,

$$\langle W_S, W_T\rangle_{L^2(\{-1,1\}^n,\nu)} = \mathbb{E}_\nu W_S(x) W_T(x) = \mathbb{E}_\nu \prod_{i\in S\Delta T} x_i = \prod_{i\in S\Delta T} \mathbb{E}_\nu x_i = \delta(T,S).$$

La base $(W_S)_{S\subset[n]}$ est appelée base de Fourier–Walsh. Ainsi, toute fonction $f \in L^2(\nu)$ se décompose en polynômes multilinéaires de manière canonique $f = \sum_{S\subset[n]} \hat{f}(S) W_s$ avec $\hat{f}(S) = \langle f, W_S\rangle_{L^2(\nu)}$.

Cette décomposition a déjà été rencontrée dans l'espace gaussien $(\mathbb{R}^n, \gamma)$. En effet les polynômes d'Hermite $h_S(x) = \prod_{i\in S} h_i(x_i)$ sont également une base orthonormée de $(\mathbb{R}^n, \gamma)$.

4.1.2 Analyse de Fourier

L'analyse de Fourier discrète comporte quelques formules connues, comme celle de Parseval :

$$\mathbb{E}_\nu f(x) g(x) = \sum_{S\subset[n]} \hat{f}(S)\hat{g}(S).$$

Ceci donne en particulier $\mathbb{E}_\nu(f) = \hat{f}(\emptyset)$, $\mathrm{Var}_\mu(f) = \sum_{S\neq\emptyset} \hat{f}^2(S)$.

L'analyse de Fourier discrète s'est avérée d'un usage récurrent pour tous les problèmes que nous citons plus bas.

4.2 Hypercontractivité du semi-groupe de Bonami–Beckner

Dans cette section, nous présentons la preuve historique du théorème d'hypercontractivité de Bonami–Beckner sur le cube discret. Ce théorème nécessite bien sûr de définir au préalable un semi-groupe sur cet espace.

4.2.1 Semi-groupe de Bonami–Beckner

Définissons L "le Laplacien discret" comme $L = \frac{1}{2}\sum_{i=1}^n D_i$ où D_i est la i-ème dérivée de $f : C_n \to \mathbb{R}$ définie par $D_i f(x) = f(\tau_i x) - f(x)$ avec $x = (x_1, \ldots, x_n)$ et $\tau_i x$ obtenu à partir de x en changeant la i-ème coordonnée. On définit la forme de Dirichlet, pour f, g des fonctions définies sur sur le cube, par

$$\mathscr{E}(f,g) = \int_{\{-1,1\}^n} f(-Lg)\,d\nu = \int_{\{-1,1\}^n} \nabla f \cdot \nabla g\, d\nu,$$

où $\nabla h = (D_1 h, \ldots, D_n h)$. Alors les éléments de la base de Fourier–Walsh sont vecteurs propres de l'opérateur L. En effet,

$$LW_S(x) = -\frac{1}{2}\sum_{i=1}^n (W_S(x) - W_S(\tau_i x)) = -|S| W_S(x)$$

car $W_S(x) - W_S(\tau_i x) = 21_{i\in S} W_S(x)$. L engendre un semigroupe $(T_t = e^{tL})_{t\ge 0}$, appelé semigroupe de Bonami–Beckner. Ainsi,

$$T_t f = \sum_{S\subset[n]} e^{-t|S|} f_S W_S = \sum_{k=0}^n e^{-tk} f^{=k},$$

où l'on note $f^{=k} = \sum_{S,|S|=k} \hat{f}(S)W_S$ la partie homogène d'ordre k. Ce semigroupe est markovien ($T_t 1 = 1$), symétrique par rapport à ν au sens que pour tout $f, g\ \{0,1\}^n \to \mathbb{R}$,

$$\int_{\{-1,1\}^n} f T_t g d\nu = \int_{\{-1,1\}^n} g T_t f d\nu,$$

les deux quantités étant égales à $\sum_{k=0}^n e^{-2tk} \langle f^{=k}, g^{=k}\rangle$.

Le semi-groupe de Bonami Beckner admet, à l'instar du semi-groupe d'Ornstein-Uhlenbeck, la représentation explicite suivante :

$$T_t f(x) = \int_{\{0,1\}^n} f(y) k_t(x,y) d\nu(y) = \int_{\{0,1\}^n} f(y) \prod_{i=1}^n (1 + e^{-t} x_i y_i) d\nu(y).$$

En effet $\prod_{i=1}^n (1 + e^{-t} x_i y_i) = \sum_{S\subset[n]} e^{-t[S|} W_S(x) W_S(y)$, de sorte que

$$\int_{\{0,1\}^n} f(y) \prod_{i=1}^n (1 + e^{-t} x_i y_i) d\nu(y) = \sum_{S\subset[n]} e^{-t[S|} \langle f, W_S\rangle W_S(x) = T_t f(x).$$

La base de Walsh s'adapte pour le cas où le cube est muni de la mesure biaisée avec un biais $p \in (0,1)$. Cette fois, on considère le cube discret comme étant $\{0,1\}^n$, qui est donc muni de $\nu_p = ((1-p)\delta_0 + p\delta_1)^{\otimes n}$. Alors $(\omega_S)_{S\subset[n]}$ où $\omega_S(x) = \prod_{i\in S} \frac{x_i - p}{\sqrt{p(1-p)}}$ est une base orthonormée de $L^2(\{0,1\}^n, \nu_p)$. On définit le “laplacien” par $L = 2p(1-p)\sum_{i=1}^n D_i$, qui vérifie de même $L\omega_S = -|S|\omega_S$. Ainsi, le semi-groupe engendré par L agit de la même façon que dans le cas uniforme :

$$T_t^p f = \sum_{S\subset[n]} e^{-t|S|} \hat{f}(S) \omega_S.$$

4.2.2 Hypercontractivité

Rappelons que la propriété d'hypercontractivité d'un semi-groupe s'écrit en général $\|P_t f\|_q \le \|f\|_r$ où $q = f(r,t) \ge r$. Dans le cas gaussien, on rappelle que $q = 1 + (r-1)e^{2t}$ est l'exposant optimal. Par commodité, on considère le cube comme étant $\{-1,1\}^n$, muni de la mesure uniforme. Sur le cube discret, par tensorisation, cette inégalité est vraie si et seulement si elle est vraie sur l'espace à deux points. En effet, avec des notations probabilistes, si $x = (\tilde{x}, x_n)$, et en supposant l'inégalité vraie sur le cube $n-1$ dimensionnel

$$\|T_t f\|_{L^q(\{-1,1\}^n)} = (\mathbb{E}_{x_n} \mathbb{E}_{\tilde{x}} (T_t f)^q)^{1/q} \le (\mathbb{E}_{x_n} (\mathbb{E}_{\tilde{x}} f^r)^{q/r}))^{1/q} \le (\mathbb{E}_{x_n} (\mathbb{E}_{\tilde{x}} f^r))^{1/r} = \|f\|_{L^r(\{-1,1\}^n)},$$

où la dernière inégalité est vraie en vertu de l'inégalité de Jensen (car $r/q \le 1$).

Sur l'espace à deux points, l'idée de Bonami et Beckner est la suivante. Par dualité, on peut supposer $1 \le r \le q \le 2$, et par homogénéité on peut considérer $f\ :\ \varepsilon \in \{-1,1\} \mapsto 1 + \varepsilon a$, avec $a \in (0,1)$. On a alors $T_t f\ :\ \varepsilon \in \{-1,1\} \mapsto 1 + \varepsilon e^{-t} a$. On cherche ainsi $q = f(r,t)$ tel que :

$$\left[\frac{1}{2}(1 - e^{-t}a)^q + \frac{1}{2}(1 + e^{-t}a)^q\right]^{1/q} \le \left[\frac{1}{2}(1-a)^r + \frac{1}{2}(1+a)^r\right]^{1/r}.$$

Comme $a \in (0,1)$, on peut directement développer en série entière, ce qui donne :

$$\left(\sum_{k\ge0} \frac{q(q-1)\cdots(q-2k+1)}{2k!} e^{-2tk} a^{2k}\right)^{r/q} \le \sum_{k\ge0} \frac{r(r-1)\cdots(r-2k+1)}{2k!} a^{2k}.$$

De plus comme $r/q \leq 1$, et par l'inégalité $(1+x)^\lambda \leq 1+\lambda x$ si $\lambda \in (0,1)$,

$$\left(\sum_{k\geq 0} \frac{q(q-1)\cdots(q-2k+1)}{2k!} e^{-2tk} a^{2k}\right)^{r/q} \leq \sum_{k\geq 0} \frac{r}{q} \frac{q(q-1)\cdots(q-2k+1)}{2k!} e^{-2tk} a^{2k}.$$

Finalement, dès que

$$(q-1)\cdots(q-2k+1)e^{-2tk} \leq (r-1)\cdots(r-2k+1),$$

l'inégalité est satisfaite.

Or $\forall k \in \mathbb{N}$, $\frac{r-k}{q-k} - \frac{r-1}{q-1} = \frac{(1-k)(q-r)}{(q-k)(q-1)} \geq 0$ si $1 \leq r \leq q \leq 2$. Ainsi pour $e^{-2t} = \frac{r-1}{q-1}$, l'inégalité est satisfaite. De plus, pour $k=1$ il y a alors égalité dans la dernière ligne de sorte que $e^{-2t} = \frac{r-1}{q-1}$ est optimal.

La contrainte $e^{-2t} \leq \frac{r-1}{q-1}$ est la même que celle sur l'espace gaussien. Ceci n'est pas innocent. En effet par le même argument de Gross que pour les modèles continus, l'hypercontractivité avec ces constantes est équivalente à l'inégalité de Sobolev logarithmique suivante :

$$\mathrm{Ent}_\nu f^2 \leq \mathscr{E}(f,f),$$

où l'on note $\mathrm{Ent}_\nu f^2 = \mathbb{E}_\nu f^2 \log f^2 - \mathbb{E}_\nu f^2 \log(\mathbb{E}_\nu f^2)$ et où $\mathscr{E}(f,f) = \sum_{i=1}^n \|D_i f\|_{L^2(\nu)}$ est la forme de Dirichlet définie plus haut.

On rappelle que l'argument de Gross consiste à étudier $F : t \mapsto \|T_t f\|_{q(t)}$, et l'on a les égalités suivantes :

$$\frac{d}{dt}(F(t)^{q(t)}) = q(t)F'(t)F(t)^{q(t)-1} + q'(t)\log F(t) F(t)^{q(t)},$$

et

$$q(t)^2 F'(t)F(t)^{q(t)-1} = q'(t)\mathrm{Ent}_\nu((P_t f)^{q(t)/2}) - q(t)^2 \mathscr{E}(P_t f, (P_t f)^{q(t)-1}).$$

Contrairement au cas continu, on n'intègre pas par parties dans la forme de Dirichlet. Pour conclure de manière analogue, on dispose du lemme suivant :

Lemme 4.2.1. $\mathscr{E}(g, g^{q-1}) \geq \frac{4(q-1)}{q^2}\mathscr{E}(g^{q/2}, g^{q/2})$.

Démonstration. Il suffit de prouver, pour tout $0 \leq a < b$,

$$\frac{4(q-1)}{q^2}(a^{q/2} - b^{q/2})^2 \leq (a-b)(a^{q-1} - b^{q-1}).$$

Or ceci résulte de l'inégalité de Cauchy–Schwarz. En effet

$$\left(\int_a^b t^{q/2-1} dt\right)^2 \leq \int_a^b dt \int_a^b t^{q-2} dt,$$

ce qui donne l'inégalité voulue.

□

Ainsi, on obtient

$$q(t)^2 F'(t)F(t)^{q(t)-1} \leq q'(t)\mathrm{Ent}_\nu((P_t f)^{q(t)/2}) - 4(q(t)-1)\mathscr{E}((P_t f)^{q(t)/2}, (P_t f)^{q(t)/2}),$$

et F est ainsi décroissante si $q(t) = 1 + (q(0)-1)e^{2\rho t}$ et ρ désigne la constante de Sobolev logarithmique.

Sur l'espace à deux points l'inégalité de Sobolev logarithmique se réécrit, en notant $u = f(1)$ et $v = f(-1)$,

$$\frac{1}{2}(u^2 \log^2 u + v^2 \log^2 v) - \frac{u^2+v^2}{2}\log(\frac{u^2+v^2}{2}) \leq \frac{(v-u)^2}{2}.$$

Cette inégalité présente l'avantage de s'établir de manière un peu plus directe que l'inégalité d'hypercontractivité de Bonami et Beckner citée plus haut. De manière plus globale, il est possible de déterminer des inégalités de Sobolev logarithmiques sur des modèles discrets plus généraux. Ceci nécessite de définir une forme de Dirichlet $\mathscr{E}$, ce que nous verrons dans la prochaine section.

En revanche une preuve d'hypercontractivité directe comme celle présentée plus haut est bien plus délicate, même sur un exemple simple comme un espace à trois points muni de la mesure uniforme.

Pour l'inégalité de Sobolev logarithmique, voyons l'exemple du cube discret muni de la mesure biaisée. La forme de Dirichlet est de manière analogue $\mathscr{E}(f,g) = \langle f, -Lg \rangle_{L^2(\nu_p)} = \langle \nabla f, \nabla g \rangle_{L^2(\nu_p)}$.

Sur l'espace à deux points, l'inégalité de Sobolev logarithmique avec constante $C(p)$ s'écrit

$$2[pu^2 \log^2 u + (1-p)v^2 \log^2 v - (pu^2 + (1-p)v^2) \log(pu^2 + (1-p)v^2)] \leq C(p)p(1-p)(v-u)^2.$$

Cette inégalité est plus difficile à établir que pour le cas uniforme. Diaconis et Saloff-Coste [D-SC] ont démontré que la meilleure constante $C(p)$ vaut $2\frac{p-(1-p)}{\log p - \log(1-p)}$. Par l'argument de Gross le semi-groupe $(T_t^p)_{t\geq 0}$ est hypercontractif avec constante $C(p)$. Notons que cette constante tend vers 0 si p tend vers 0.

4.3 Généralisation à des espaces produits et aux graphes de Cayley/Schreier

Soit Ω un espace fini muni d'une mesure de probabilité μ avec une chaine de Markov K, invariante et réversible par rapport à μ, i.e. telle que :

$$\forall (x,y) \in \Omega^2, \quad \sum_{x \in \Omega} K(x,y)\mu(x) = \mu(y) \quad \text{et} \quad K(x,y)\mu(x) = K(y,x)\mu(y).$$

Soit $L = K - Id$, générateur du semi-groupe $P_t = e^{tL}$, $t \geq 0$.

La forme de Dirichlet associée est la forme bilinéaire symétrique suivante :

$$\mathscr{E}(f,g) = \int_\Omega f(-Lg)d\mu = \frac{1}{2} \sum_{x,y \in \Omega} (f(x)-f(y))(g(x)-g(y))K(x,y)\mu(y)$$

pour f, g des fonctions définies sur Ω.

Un exemple intéressant qui englobe le cas du cube discret est donné par un produit de tels espaces munis de la mesure produit suivante : $\Omega = \Omega_1 \times \cdots \times \Omega_n \quad \text{avec} \quad \mu = \mu_1 \otimes \cdots \otimes \mu_n$, où l'on prend la chaîne de Markov produit. C'est-à-dire, si $i = 1, \ldots, n$, et $f : \Omega \to \mathbb{R}$ $L_i f = \int_{\Omega_i} f d\mu_i - f$ on considère le générateur sur l'espace produit

$$Lf = \sum_{i=1}^n L_i f.$$

La forme de Dirichlet $\mathscr{E}$ se décompose ainsi en

$$\mathscr{E}(f,f) = \sum_{i=1}^n \int_{\Omega_i} L_i(f)^2 d\mu_i. \tag{4.1}$$

Ce cadre, que l'on appellera cadre produit contient le cube biaisé avec $\Omega_1 = \cdots = \Omega_n = \{-1, 1\}$, où chacun des espaces est équipé de la mesure $p\delta_{-1} + q\delta_1$.

Parmi les exemples classiques, cela recouvre également les tores discrets, c'est-à-dire les exemples où pour tout i, $\Omega_i = \mathbb{Z}/m\mathbb{Z}$ pour $m \geq 3$ muni de la mesure uniforme. Ces modèles produits sont un cas particulier de graphes de Schreier et de Cayley. Des études récentes ont été menées sur de tels graphes ([CE-L], [O-W]), qui contiennent des exemples non produits. Nous définissons maintenant ces graphes.

Soit G un groupe fini contenant un sous-ensemble générateur S, stable par conjugaison et agissant transitivement sur un ensemble fini Ω. Pour $g \in G$, $x \in \Omega$, soit x^g l'action de g sur x. On définit alors le graphe de Schreier ayant pour sommets Ω et pour arêtes (x, y) lorsqu'il existe un $s \in S$ tel que $y = x^s$.

Un graphe de Cayley est le cas particulier où le groupe agit sur lui même, c'est-à-dire $G = \Omega$. Dans ce cas, le graphe (G, V) est donc l'ensemble des sommets $(g)_{g\in G}$ avec arêtes $(g, gs)_{g\in G, s\in S}$. Pour les applications qui suivent, nous ferons l'hypothèse que le sous-ensemble générateur S est symétrique, c'est-à-dire $S^{-1} = S$.

Deux exemples non produits importants de tels graphes sont donnés par les tranches du cube booléen (ou modèle de Bernoulli–Laplace) et le groupe symétrique $\mathfrak{S}_n$.

Le sous-ensemble des transpositions $\mathscr{T}_n$ engendre le groupe symétrique $\mathfrak{S}_n$ et satisfait l'hypothèse de symétrie. Ainsi le graphe de Cayley correspondant est donné par $(\sigma, \sigma\tau)_{\sigma\in\mathfrak{S}_n, \tau\in\mathscr{T}_n}$.

Pour $n \geq 1$, $1 \leq k < n$, les tranches d'ordre k du cube booléen sont définies par

$$\binom{[n]}{k} = \{x \in \{0,1\}^n, \sum_{i=1}^{n} x_i = k\}.$$

Le groupe agissant transitivement sur $\binom{[n]}{k}$ est le groupe symétrique $\mathfrak{S}_n$ par $x^\sigma = (x_{\sigma(i)})_{1\leq i\leq n}$, et le graphe associé est $(x, x^{(ij)})_{\{x\in\binom{[n]}{k}, (ij)\in\mathscr{T}_n\}}$.

Étant donné un graphe de Schreier ou de Cayley G, soit μ la mesure uniforme et K le noyau de transition donné par

$$K(g_1, g_2) = \frac{1}{|S|} 1_S(g_1 g_2^{-1}), \quad g_1, g_2 \in G,$$

qui correspond à la marche aléatoire au plus proche voisin.

Ce noyau engendre un semi-groupe $(P_t = e^{tL})_{t\geq 0}$, où $L = K - Id$. La forme de Dirichlet associée $\mathscr{E}$ s'écrit ainsi :

$$\mathscr{E}(f, f) = \frac{1}{2|S|} \sum_{g\in G} \sum_{s\in S} [f(gs) - f(g)]^2 \mu(g) = \frac{1}{2|S|} \sum_{s\in S} \|D_s f\|^2_{L^2(G)},$$

où l'on note $D_s f : g \mapsto f(gs) - f(g)$ dans le cas d'un graphe de Cayley et

$$\mathscr{E}(f, f) = \frac{1}{2|S|} \sum_{x\in\Omega} \sum_{s\in S} [f(x^s) - f(x)]^2 \mu(x) = \frac{1}{2|S|} \sum_{s\in S} \|D_s f\|^2_{L^2(\Omega)},$$

où $D_s f : x \mapsto f(x^s) - f(x)$ dans le cas d'un graphe de Schreier.

Par exemple dans le cas des tranches du cube ou du groupe symétrique, avec la normalisation ci-dessus, $Lf = \frac{2}{n(n-1)} \sum_{1\leq i<j\leq n} D_{\tau_{ij}} f$. De plus, on a alors $P_t f(x) = e^{-t} \sum_{m\geq 0} \frac{t^m}{m!} K^m f(x) = \mathbb{E}_\mu f(y)$ où y est obtenu à partir de x en effectuant m transpositions aléatoires, où $m \sim \mathscr{P}(t)$ (loi de Poisson de paramètre t).

Dans ce contexte, on dit que (Ω, L, μ) satisfait à une inégalité de trou spectral $SG(\kappa)$ si :

$$\forall f : \Omega \to \mathbb{R}, \ \kappa \mathrm{Var}_\mu(f) \leq \mathscr{E}(f, f). \tag{4.2}$$

De même, on dit que (Ω, Γ, μ) satisfait à une inégalité de Sobolev logarithmique $LS(\kappa)$ si :

$$\forall f : \Omega \to \mathbb{R}, \ \kappa \, \mathrm{Ent}_\mu(f^2) \leq 2\,\mathscr{E}(f, f). \tag{4.3}$$

Nous noterons λ (λ_n) et ρ (ρ_n) les meilleures (i.e. plus grandes) constantes dans ces inégalités. De façon générale on a toujours $\lambda \geq \rho$ (par le même argument de linéarisation qu'en continu).

Dans la cadre produit $\Omega = \Omega_1 \times \cdots \times \Omega_N$, ces inégalités se tensorisant, on a $\lambda = \min(\lambda_1, \ldots, \lambda_N)$ et $\rho = \min(\rho_1, \ldots, \rho_N)$.

Notons qu'avec les normalisations choisies, il y a un facteur $\frac{1}{n}$ dans ces constantes par rapport à celles données sur le cube booléen. L'argument de Gross implique de nouveau de manière analogue l'hypercontractivité, qui est un outil fondamental pour de nombreuses applications, pour des modèles discrets satisfaisant une inégalité de Sobolev logarithmique. Ainsi, les inégalités de Sobolev logarithmique ont été très étudiées dans ce cadre. Une référence fournie à ce sujet est l'article de Diaconis et Saloff-Coste [D-SC]. Donnons quelques exemples explicites.

Pour les tores discrets $(\mathbb{Z}/m\mathbb{Z})^n$, la constante de trou spectral vaut, dans la normalisation des graphes de Cayley, $\lambda = \frac{1-\cos(\frac{2\pi}{m})}{n}$, celle de Sobolev logarithmique est égal à celle du trou spectral pour m impair, et déterminée pour m pair par Chen et Sheu [Ch-Sh]. Dans les deux cas, ces constantes sont de l'ordre de $\frac{1}{nm^2}$ pou m grand. Pour les groupes symétriques $\mathfrak{S}_n$, , on a $\lambda_n = \frac{1}{n}$ et $\rho_n = \frac{a}{n\log n}$ pour $a > 0$ [D-SC]. Sur les tranches du cube d'ordre k, $\lambda_n = \frac{1}{n}$, et ρ_n satisfait $\rho_n^{-1} \sim n \log \frac{n^2}{k(n-k)}$. Ce résultat a été obtenu par Lee et Yau [L-Y]. Il est intéressant de constater que lorsque k est de l'ordre de n, la constante de Sobolev logarithmique est du même ordre que celle du trou spectral - ce qui n'est pas le cas pour les groupes symétriques. Ceci est à l'origine du fait que la plupart des résultats présentés dans la section suivante s'étendent dans le cas des tranches du cube et non dans le cas des groupes symétriques. Nous discuterons ce point plus en détails dans la présentation de nos résultats.

4.4 Influences

Cette section est destinée à présenter une notion fondamentale dans l'analyse des fonctions booléennes, la notion d'influence. Nous avons déjà défini la notion d'influence géométrique dans le cas continu, directement inspirée par celle plus ancienne d'influence d'une fonction booléenne. Sur le cube, la notion d'influence est claire et possède une définition unique.

Si l'on pense à $f : \{-1,1\}^n \mapsto \{-1,1\}$ comme une fonction de choix pour un vote, l'influence de la i-ème coordonnée mesure la probabilité que le résultat de l'élection change si le i-ème votant change son vote. Cette définition s'étend à des fonctions à valeurs réelles, comme nous le verrons plus bas. Dans la suite, on considérera des fonctions à valeurs dans $\{0,1\}$ (ou $\{-1,1\}$), $[0,1]$ ou $\mathbb{R}$.

Définition 4.4.1. *Soit $f : \{0,1\}^n \mapsto \{0,1\}$. La i-ème coordonnée est dire "pivotale" si $f(x) \neq f(x^{\oplus i})$ où $x^{\oplus i}$ désigne le vecteur obtenu à partir de x en changeant le signe de la i-ème coordonnée.*

Définition 4.4.2. *Soit $f : \{0,1\}^n \mapsto \{0,1\}$. L'influence de la i-eme coordonnée est donnée par*

$$I_i(f) = \mathbb{P}_{x\sim\nu}(f(x) \neq f(x^{\oplus i}))$$

où $x^{\oplus i}$ désigne le vecteur obtenu à partir de x en changeant le signe de la i-ème coordonnée.

La somme de toutes les influences est appelée somme totale (également nommée initialement "average sensitivity"), et est souvent notée $I(f)$. Lorsque f est la fonction caractéristique d'un sous-ensemble du cube A, la somme totale représente la fraction d'arêtes reliant un point de A à un point de $\{-1,1\}^n \backslash A$. C'est ce qu'on définit comme la mesure de bord d'un ensemble, et ainsi la somme des influences est liée à un problème isopérimétrique discret.

Notons que pour des fonctions à valeurs dans $\{0,1\}$, (à une normalisation près) $I_i(f) = \|D_i f\|_1 = \|D_i f\|_p^p$ où D_i est l'opérateur de dérivation discrète défini plus haut. En lien avec l'analyse de Fourier discrète, il est intéressant d'exprimer les influences en termes spectraux. Il n'est pas difficile de voir que

$$I_i(f) = \sum_{i\in S} \hat{f}(S)^2.$$

Cette remarque étant faite, on définit plus généralement l'influence $I_i(f)$ de la i-ème coordonnée d'une fonction $f : \{0,1\}^n \to \mathbb{R}$ par

$$\mathbb{E}_{x_i}(\mathrm{Var} f) = \sum_{i\in S} \hat{f}(S)^2.$$

En conséquence, la somme totale d'influences s'écrit de manière spectrale

$$I(f)=\sum_{i=1}^{n} I_i(f)=\sum_{S\subset[n]}|S|\hat{f}(S)^2=\sum_{k\geq 0}k\|f^{=k}\|_2^2.$$

Parmi les exemples les plus célèbres, les fonctions Majorité et Dictateurs sont définies par

$$\mathrm{Maj_n}\{0,1\}^n \;\rightarrow\; \{0,1\},\, x \;\mapsto\; \mathbf{1}_{\{\sum_{i=1}^n x_i\geq n/2\}} \quad \text{et} \quad \mathrm{Dic}_i\{0,1\}^n\rightarrow\{0,1\},\, x\mapsto x_i.$$

Ces deux fonctions seront dans la suite des fonctions modèles.

Alors il est immédiat que $I_j(\mathrm{Dic}_i)=\delta_{ij}$. Il n'est pas difficile de démontrer que les influences de la fonction majorité sont de l'ordre de $\frac{1}{\sqrt{n}}$, i.e. pour tout $i\in\{1,\dots,n\}$, $\mathrm{I}_i(\mathrm{Maj_n})\sim\frac{1}{\sqrt{n}}$.

Sur les exemples de modèles discrets présentés dans la section précédente, la notion d'influence s'étend naturellement à partir de la remarque $I_i(f)=\|D_i f\|_1$. Sur un graphe de Schreier ou de Cayley - cela inclut le cadre produit, nous avons défini un opérateur de dérivation discrète, L_i ou D_s. La définition d'influence s'étend ainsi naturellement.

Définition 4.4.3. *L'influence de la i-ème coordonnée $f\,:\,\Omega_1\times\cdots\times\Omega_n\rightarrow\mathbb{R}$ est donnée par*

$$I_i(f)=\|L_i f\|_1.$$

Il est à noter que sur le cube discret, avec les notations précédentes, il y a une différence de normalisation entre les deux définitions. Pour un graphe de Schreier ou de Cayley quelconque, on a la définition suivante :

Définition 4.4.4. *L'influence d'un élément $s\in S$ pour $f\,:\,G\rightarrow\mathbb{R}$ est donnée par $I_s(f)=\|D_s f\|_1$.*

Dans le cas des tranches du cube (d'ordre k), les influences des transpositions $I^k_{\tau_{ij}}f$ sont reliées aux influences des coordonnées $I_i f$ sur le cube. Définissons l'influence d'une transposition sur le cube par $I_{\tau_{ij}}f=\mathbb{P}(f(x^{(ij)})\neq f(x))$. Alors, pour tout i, $I_i(f)=\frac{1}{n}\sum_{j=1}^n I_{\tau_{ij}}f$, et de plus, par formule des probabilités totales, pour toute transposition τ_{ij},

$$I_{\tau_{ij}}f=\sum_{k=0}^{n}\mathbb{P}(x_1+\cdots+x_n=k)I^k_{\tau_{ij}}f=\frac{1}{2^n}\sum_{k=0}^{n}\binom{k}{n}I^k_{\tau_{ij}}f.$$

Par exemple, si toutes les influences de f tendent vers 0, les influences sur les tranches tendent également vers 0 pour toutes les tranches sauf une proportion négligeable.

Plus bas, nous mentionnerons un travail de Filmus [Fil] qui permet de relever sur le cube des fonctions définies seulement sur des tranches de celui-ci. Là encore, les influences correspondantes sont liées.

4.5 Principaux résultats sur les influences

4.5.1 Théorème isopérimétrique de Harper

Par convention, ici, le cube discret sera $\{0,1\}^n$. On le munit de l'ordre lexicographique. On définit la distance de Hamming

$$d(x,y)=\sum_{i=1}^{n}|x_i-y_i|=\sharp\{i\in\{1,\dots,n\},\,x_i\neq y_i\}.$$

Une arête est un couple (xy) avec $(x,y)\in\{0,1\}^n\times\{0,1\}^n$ tel que $d(x,y)=1$ (x et y sont dits "voisins").

Le problème isopérimétrique sur le cube discret nécessite une définition du bord. On définit le bord arête par $\partial_e A = \{(xy),\ x \in A,\ y \notin A\}$. L'inégalité isopérimétrique de Harper [Har] donne ainsi une borne inférieure sur la somme totale des influences. Les ensembles extrémaux sont des réunions de sous-cubes, c'est-à-dire les sous-ensembles d'ordre m constitués des m premiers sommets pour l'ordre lexicographique.

Théorème 4.5.1. *Soit $A \subset \{0,1\}^n$ tel que $|A| = m \leq 2^{n-1}$, et C un sous cube d'ordre m. Alors*

$$\frac{|\partial_e A|}{|A|} \geq \frac{|\partial_e C|}{|C|} = m(n - \log_2 m)$$

En termes d'influences, ce résultat s'énonce de la façon suivante :

$$I(f) \geq -\mathbb{E}_\nu f \log_2 \mathbb{E}_\nu f.$$

En conséquence, il existe une influence au moins de l'ordre de $\frac{1}{n}$. Ce résultat se retrouve en utilisant l'inégalité d'Efron Stein - ou de Poincaré. En effet, ce résultat se réécrit :

$$\mathrm{Var}_\mu(f) \leq \frac{1}{4} I(f).$$

4.5.2 Conjecture Entropie/Influence

L'inégalité de Sobolev logarithmique donne une autre borne inférieure pour la somme totale des influences, c'est-à-dire

$$\mathrm{Ent}_\nu(f^2) \leq C I(f),$$

où C est une constante numérique qui ne dépend que de la normalisation choisie. Au milieu des années 90, dans leur étude sur les phénomènes de seuil, Friedgut et Kalai [F-K] conjecturent une version spectrale de cette inégalité :

Conjecture 4.5.2. *Pour toute fonction $f\ :\ \{-1,1\}^n \to \{-1,1\}$, il existe une constante C universelle telle que :*

$$\sum_{S \subset [n]} \hat{f}^2(S) \log\left(\frac{1}{\hat{f}^2(S)}\right) \leq C I(f) = C \sum_{S \subset [n]} |S| \hat{f}(S)^2.$$

La condition d'être booléenne est nécessaire. Cette conjecture est intéressante car elle implique de nombreux résultats intéressants. En particulier, elle implique des phénomènes de seuils étroits. Elle contient également le théorème de Kahn-Kalai-Linial que nous énonçons plus bas. Elle permettrait de mieux relier le comportement du spectre de la fonction f en fonction de la somme de ses influences. Enfin, elle est intimement liée à des problèmes d'informatique théorique.

Elle est appelée conjecture entropie/influence, et c'est encore un problème ouvert aujourd'hui. Seuls quelques cas particuliers de cette conjecture ont été résolus (par exemple par O'Donnell et Zhou [OD-Z] en faisant l'hypothèse que f est symétrique, c'est-à-dire invariante par permutation des coordonnées).

4.5.3 Tribes, KKL, Talagrand

Concernant le comportement du maximum des influences, un résultat pionnier est dû à Ben-Or et Linial ([BO-L]). De manière surprenante, les auteurs exhibent des fonctions booléennes, appelés *tribes* telles que l'influence de chaque coordonnée n'excède pas $\frac{\log n}{n}$. La fonction majorité n'est donc pas celle qui minimise le maximum des influences, car ces dernières sont toutes de l'ordre de $n^{-1/2}$. Ces fonctions sont définies de la façon suivante. Si $n = km$,

$$\mathrm{Tribes}_{km} x = ((x_1, \ldots, x_k), \ldots, (x_{m(k-1)+1}, \ldots, x_{km})) \in \{0,1\}^{km}$$

vaut 1 si et seulement si l'une des tribus de longueur k $(x_{j(k-1)+1}, \ldots, x_{jk})$ est la tribu où tous les x_i sont égaux à 1. Ben-Or et Linial démontrent que pour n grand, si $k \approx \log n - \log\log n$, toutes les influences sont au plus de l'ordre de $\frac{\log n}{n}$, plus précisément la proposition suivante :

Proposition 4.5.3.

$$I_i(\mathrm{Tribes}_n) = \frac{\log n}{n}\Big(1+o(1)\Big).$$

Démonstration. Premièrement la probabilité qu'une des tribus est la tribu $(1,\ldots,1)$ vaut 2^{-k}. Ainsi, par borne d'union, $\mathbb{P}(\mathrm{Tribes}_n = 0) = (1-2^{-k})^m$. Ensuite i est pivotale si et seulement si tous les éléments de la tribu contenant i sont tous égaux à 1 et qu'aucune autre tribu n'est dans cette configuration. Ceci arrive avec probabilité $2^{-(k-1)}(1-2^{-k})^{m-1}$ - ce qui est par définition l'influence de la i-ème coordonnée. Cette quantité est minimale si $k = \log n - \log\log n + \log\log 2$, et vaut alors $\frac{\log n}{n}\Big(1+o(1)\Big)$. □

Ben Or et Linial conjecturent que ce résultat est optimal. Motivés par ce résultat, Kahn, Kalai et Linial démontrent la conjecture de Ben Or et Linial au sens où toute fonction booléenne possède une influence plus grande que $\frac{\log n}{n}$. C'est le théorème KKL, qui est un résultat fondamental dans l'analyse des fonctions booléennes, où l'analyse de Fourier discrète en lien avec l'hypercontractivité du semi-groupe de Bonami-Beckner joue un rôle crucial dans la preuve.

Théorème 4.5.4. *Soit f $\{0,1\}^n \to \{0,1\}$. Alors il existe $i \in \{1,\ldots,n\}$, $c>0$,*

$$I_i(f) \geq c\mathrm{Var}_\nu(f)\frac{\log n}{n}.$$

Peu après, ce résultat sera étendu au cube muni de la mesure biaisée de paramètre p indépendant de n [BKal]. Inspiré par ces travaux, Talagrand [Tal1] démontre une inégalité sur le cube biaisée de paramètre p, appelée désormais inégalité de Talagrand qui implique le résultat de Kahn-Kalai-Linial et sa généralisation au cas biaisé. Là encore les outils utilisés dans la preuve de ce résultat font appel à l'analyse de Fourier discrète associée avec l'hypercontractivité.

Théorème 4.5.5. *Soit f $\{0,1\}^n \to \{0,1\}$. Alors*

$$\mathrm{Var}_{\nu_p}(f) \leq C(p)\sum_{i=1}^n \frac{I_i(f)}{1+\log(I_i(f))}.$$

Il n'est pas difficile de vérifier que cette inégalité implique bien l'existence d'une coordonnée i telle que $I_i(f) \geq C'(p)\mathrm{Var}_{\nu_p}(f)\frac{\log n}{n}$. Ce n'est que récemment ([O-W], [CE-L]) que ces résultats ont été étendus à des espaces discrets plus généraux tels que les graphes de Schreier et de Cayley dont on connait une borne explicite sur la constante d'hypercontractivité. Ces exemples recouvrent les tores discrets $(\mathbb{Z}/m\mathbb{Z})^n$, le modèle de Bernouilli-Laplace $\binom{[n]}{k}$ ou les groupes symétriques $\mathfrak{S}_n$.

En fait, l'inégalité démontrée par Talagrand est valable pour des fonctions à valeurs réelles et s'écrit, pour $f : \{0,1\}^n \to \mathbb{R}$:

$$\mathrm{Var}_{\nu_p}(f) \leq C(p)\sum_{i=1}^n \frac{\|D_i f\|_2^2}{1+\log(\|D_i f\|_2/2\sqrt{p(1-p)}\|D_i f\|_1)}.$$

Cette inégalité est une version améliorée de l'inégalité de Poincaré. Nous avons vu dans le cadre continu que la théorie des semi-groupes est un outil très efficace vis-à-vis de cette inégalité. En utilisant la décomposition de la variance le long du semi-groupe de la chaleur et en utilisant l'hypercontractivité du semi-groupe, on retrouve l'inégalité de Talagrand (cf [CE-L]). Ce schéma de preuve de l'inégalité de Talagrand s'adapte en fait très bien sur des modèles où la forme de Dirichlet se décompose le long de "directions".

4.5.4 Le théorème “Junta” de Friedgut

Parmi les autres résultats, un théorème fondateur est celui de Friedgut [Fri], le théorème “Junta”. Une *junta*, junte en francais est une fonction ne dépendant que d’un nombre fixé de coordonnées. Si f est une junte, c’est une conséquence immédiate que sa somme totale d’influences est bornée. Le théorème de Friedgut s’énonce de la façon suivante :

Théorème 4.5.6. *Soit $f : \{0,1\}^n \mapsto \{0,1\}$ d’influence totale égale à I. Alors f est ε proche d’une $e^{O(I/\varepsilon)}$-junta, c’est-à-dire d’une fonction dépendant au plus de $e^{O(\mathbb{I}/\varepsilon)}$-coordonnées.*

Ce théorème trouve son intérêt si l’on pense à la somme totale des influences comme le nombre de coupures dans un graphe. Ce n’est que récemment que ce théorème a été étendu de façon plus générale sur des graphes de Cayley produits ([Sa-Tu]) ou sur le tore continu [Aus]. Un résultat analogue, suite à la définition d’influences géométrique, apparait dans [K-M-S1] pour certaines classes de mesures de Bolzmann. La question de l’étendre sur des exemples de graphes non produits reste en grande partie ouverte. En effet, les principales preuves utilisent, soit l’analyse de Fourier discrète ([Sa-Tu]), soit de façon cruciale la structure produit du graphe ([Aus]). Une seule généralisation sur l’exemple du modèle de Bernouilli–Laplace a été étendue par Wimmer [Wim] puis indépendamment par Filmus [Fil], là encore en utilisant des techniques d’analyse de Fourier discrète. Dans le dernier chapitre, nous donnerons un schéma de preuve commun basé sur l’argument d’Austin et qui a l’avantage de démontrer un théorème analogue à celui de Friedgut pour tous les exemples produits. Les constantes sont toutefois moins bonnes que celles précédemment obtenues.

4.5.5 Théorèmes de structures

Pour clore cette section, mentionnons un autre résultat ne faisant pas intervenir la notion d’influence mais uniquement les coefficients de Fourier : le théorème de Friedgut–Kalai–Naor [F-K-N]. Il établit que si les coefficients de Fourier d’une fonction booléenne f sont essentiellement concentrés sur le premier niveau, f est proche d’une fonction dictateur.

Théorème 4.5.7. *Soit $f^{>k} = \sum_{k<|S|\leq n} \hat{f}(S) W_S$. Si $\|f^{>1}\|_2^2 \leq \varepsilon$, il existe $i \in [n]$ tel que $\|f - \mathrm{Dic}_i\|_2 = O(\varepsilon)$.*

La généralisation à des niveaux supérieurs de ce résultat est dû à Kindler et Safra [Ki-Sa] :

Théorème 4.5.8. *Si $\|f^{>k}\|_2^2 \leq \varepsilon$, il existe une fonction h dépendant de $O(k,p)$ coordonnées de degré au plus k tel que $\|f - h\|_2 = O_{k,p}(\varepsilon)$.*

Les deux résultats précédents s’énonçant en termes de décomposition spectrale, il est plus difficile a priori de les étendre à des exemples de graphes non produits. La construction de Filmus permet toutefois de les étendre sur le modèle de Bernouilli Laplace. Dans [F-K-M-W], les auteurs se servent de ce résultat pour des problèmes de combinatoire du type Erdös–Ko–Rado sur l’intersection de familles d’ensembles à t éléments.

4.6 La notion de noise stability/sensitivity

Repensons à une fonction booléenne sur le cube discret comme une fonction de choix pour le résultat d’une élection. Aucune élection n’étant infaillible, on peut se demander ce qu’il se passe s’il y a une probabilité (petite) que chacun des votes soit mal interprété. De la même façon, dans d’autres domaines, on s’intéresse à des petites perturbations (bruits) de la distribution initiale donnée sur le cube booléen. C’est pourquoi on introduit naturellement la notion de “noise stability” booléenne, de la même façon que sur l’espace gaussien.

En théorie de la percolation, Benjamini Kalai et Schramm ont popularisé cette notion. Dans [B-K-S], les auteurs démontrent entre autres que la probabilité d’avoir un chemin de gauche à droite

(une percolation) dans le rectangle $[n] \times [n+1] \subset \mathbb{Z}^2$ est asymptotiquement sensible au bruit au sens défini plus bas. Sans rentrer dans les détails, cela veut dire la chose suivante.
Considérons une configuration où il y a une percolation de gauche à droite : si l'on lui ajoute un bruit η indépendant de n (même très petit), alors il n'est plus possible de prédire s'il y a encore percolation. Autrement dit, si f est la fonction booléenne correspondante à l'événement de percolation et X est un vecteur aléatoire de $\{0,1\}^n$, la connaissance de $f(X)$ ne donne pas d'information sur $f((1-\eta)X+\eta X')$, où X' est un autre vecteur aléatoire de $\{0,1\}^n$ indépendant de X.

4.6.1 Le cas du cube discret

Nous avons déjà défini la notion de stabilité au bruit dans le cadre gaussien. Rappelons que dans ce cadre, $\mathscr{S}_\eta(f) = \mathbb{E}_\gamma[f(X)f(Y)]$ où X et Y sont des variables gaussiennes η-corrélées. La définition est identique sur le cube.

Définition 4.6.1.
$$\mathscr{S}_{\eta,b}(f) = \mathbb{E}_\nu[f(X)f(X^\eta)] = \mathbb{E}_{\nu\otimes\nu}[f(X)f(Y)],$$
où X, Y sont des variables distribuées selon ν et η-corrélées.

La notion de stabilité au bruit est reliée à celle duale de sensibilité au bruit (“noise sensitivity” en anglais). La sensibilité au bruit d'une fonction est la quantité $\mathscr{N}\mathscr{S}_{\eta,b}(f) = \mathscr{S}_{\eta,b}(f) - \mathbb{E}_\nu^2(f)$. Une suite de fonctions booléennes est dite asymptotiquement sensible au bruit, ou simplement sensible au bruit, si la limite $\lim_{n\to\infty} \mathscr{N}\mathscr{S}_{\eta,b}(f_n)$ converge vers 0. De même la stabilité au bruit d'un ensemble est défini par la stabilité au bruit de sa fonction caractéristique.

Le principal résultat de l'article [B-K-S] est de prouver un critère de sensibilité au bruit pour des fonctions booléennes. Ce critère est le suivant : si la somme des carrés des influences d'une suite de fonctions booléennes $(f_n)_{n\geq 0}$ tend vers 0, alors la suite $(f_n)_{n\geq 0}$ est sensible au bruit. Comme application, la fonction de percolation de premier passage dans un rectangle est sensible au bruit. En fait les auteurs donnent plusieurs exemples de fonctions booléennes issues de modèles de percolation où les influences sont calculables et le critère utilisable.

Comme pour l'espace gaussien, la sensibilité au bruit peut être exprimée en terme de semi-groupe donc de manière spectrale de la façon suivante. Soit $\eta = e^{-t}$,
$$\mathscr{S}_{\eta,b}(f) = \mathbb{E}_\nu(f T_t f) = \sum_{k=0}^{\deg f} \eta^k \|f^{=k}\|_2^2,$$
et de même pour le cube biaisé,
$$\mathscr{S}_\eta(f) = \mathbb{E}_{\nu_p}(f T_t^p f) = \mathbb{E}_{\nu_p}((T_{t/2}^p f)^2).$$
La sensibilité au bruit, pour un bruit η, correspond ainsi à la covariance entre f et $T_{-\log\eta} f$. Parmi les questions naturelles, on peut se demander quelles sont les fonctions les plus stables au bruit ou au contraire celles qui sont sensibles au bruit.

Sans hypothèses spécifiques, les fonctions booléennes non constantes de moyenne nulle les plus stables au bruit sont les fonctions dictateurs. Ce résultat est immédiat en utilisant la décomposition de Fourier plus haut. En effet, si f de moyenne nulle, $\sum_{k=0}^{\deg f} \eta^k \|f^{=k}\|_2^2 \leq \eta$ avec égalité si et seulement pour tout $k \neq 1$, $\|f^{=k}\|_2 = 0$, c'est-à-dire si f est un dictateur.

En fait, l'idée de départ des spécialistes est qu'étudier le spectre de la fonction f donne des informations sur sa sensibilité au bruit. Intuitivement, plus le spectre est concentré sur les premiers niveaux, plus la fonction est stable au bruit. Au contraire, une fonction sensible au bruit devrait avoir un spectre très étendu. Rappelons que la formule de Parseval implique que $\sum_{i=1}^n \|f_n^{=k}\|_2^2 = 1$. De plus, si η est fixé, il est clair que
$$\lim_{n\to\infty} \mathscr{N}\mathscr{S}_\eta(f_n) = 0 \iff \forall m \geq 0,\ \lim_{n\to\infty} \sum_{k=1}^m \|f_n^{=k}\|_2^2 = 0.$$

Cette équivalence est à l'origine du travail de Benjamini, Kalai et Schramm [B-K-S]. En adaptant une inégalité de Talagrand, les auteurs bornent les coefficients $\|f_n^{=k}\|_2^2$ en termes de la somme des carrés des influences $\sum_{i=1}^n I_i^2(f_n)$. De là ils en tirent le critère suivant, dit de Benjamini–Kalai–Schramm :

Théorème 4.6.2. *Soit $(f_n)_{n\geq 0}$ une suite de fonctions booléennes définies sur $\{0,1\}^n$. Alors :*

$$\lim_{n\to\infty}\sum_{i=1}^n I_i^2(f_n)=0 \quad\Longrightarrow\quad \lim_{n\to\infty}\mathcal{N}\mathcal{S}_\eta(f_n)=0.$$

4.6.2 Espaces discrets et graphes de Schreier

De manière plus générale, à la lumière des exemples booléens et gaussiens, on peut définir une notion de "noise stability" sur les espaces discrets présentés plus haut en utilisant l'expression $\mathbb{E}_\mu((P_{t/2}f)^2)$. Toutefois, nous avons défini le bruit η comme une quantité entre 0 et 1. Il reste donc à déterminer la relation entre le bruit η et le paramètre du semi-groupe t. L'inégalité de trou spectral suivante nous éclaire sur ce point : si f est de moyenne nulle,

$$\mathbb{E}_\mu((P_tf)^2)\leq e^{-\lambda t}\mathbb{E}_\mu(f^2).$$

L'inégalité de trou spectral implique ainsi une convergence exponentielle vers l'équilibre proportionnelle au trou spectral λ.

Définition 4.6.3. *Posons $\eta=e^{-\lambda t}$. On définit la "noise stability" d'une fonction $f:G\to\mathbb{R}$ par*

$$\mathcal{S}_\eta(f)=\mathbb{E}_\mu((P_{t/2}f)^2)=\mathbb{E}_{\mu\otimes\mu}(f(x)f(y)),$$

où y est obtenu en faisant agir sur x aléatoirement m éléments de S où $m\sim\mathcal{P}(\lambda^{-1}\log(\eta^{-1}))$.

On définit de même la "noise sensitivity" par $\mathcal{N}\mathcal{S}_\eta(f)=\mathcal{S}_\eta(f)-(\mathbb{E}_\mu f)^2$. Avec cette définition, la majoration donnée par l'inégalité de trou spectral est similaire à celle du cube : $\mathcal{N}\mathcal{S}_\eta(f)\leq\eta\mathrm{Var}_\mu(f)$. En particulier, si η est fixé et $(\|f_n\|_2)_{n\geq 0}$ uniformément minorée, l'inégalité de trou spectral ne donne pas d'information sur le comportement de la noise sensitivity $\mathcal{N}\mathcal{S}_\eta(f_n)$ de la suite $(f_n)_{n\geq 0}$.

Parmi les questions naturelles, se posent les suivantes : étant donné une suite de fonctions $(f_n)_{n\geq 0}$ sur des graphes $(\Omega_n)_{n\geq 0}$ (où implicitement le cardinal de Ω_n tend vers l'infini), sous quelles conditions $(f_n)_{n\geq 1}$ est sensible au bruit (i.e. $\mathcal{S}_\eta(f_n)$ converge vers 0) ? Au contraire y-a-t-il une fonction de η qui borne les suites $(\mathcal{S}_\eta(f_n))_{n\geq 0}$?

Concernant la "noise sensitivity", nous répondrons partiellement à la question en étendant le critère de Benjamini–Kalai–Schramm en donnant de plus une version quantitative. En particulier, qualitativement, nous obtenons le critère suivant :

Théorème 4.6.4. *Soit $(G_n)_{n\geq 0}$ une suite de graphes de Schreier engendrés par S_n, de constantes de trou spectral λ_n et de Sobolev logarithmique ρ_n. Supposons que le rapport $\frac{\rho_n}{\lambda_n}$ soit uniformément minoré par $c>0$. Soit $(f_n)_{n\geq 0}$ une suite de fonctions de $G_n\to\mathbb{R}$ telle que $\sum_{s\in S_n}I_s(f_n)^2\to 0$ quand $n\to\infty$. Alors $(f_n)_{n\geq 0}$ est sensible au bruit, c'est-à-dire que pour tout $\eta>0$, $\mathcal{S}_\eta(f_n)\to 0$ quand $n\to\infty$.*

Ce résultat sera présenté dans le chapitre suivant et est un cas particulier des travaux présentés dans [Bou].

Remarquons que ce théorème n'est pas vrai quand $\frac{\rho_n}{\lambda_n}$ tend vers 0. Un exemple est donné par le cube discret $\{0,1\}^N$ muni de la mesure biaisée avec un biais de $1/N$.
Soit $N=\frac{n(n-1)}{2}$, et un graphe à n sommets sur lequel une arrête est mise avec probabilité $1/N$. Si f_n désigne la fonction booléenne qui vaut 1 si et seulement le graphe contient un triangle, alors $\sum_{i=1}^N I_i^2(f_n)$ tend vers 0 tandis que que $(f_n)_{n\geq 0}$ est stable au bruit.

Pour ce qui est du problème de maximiser la “noise stability”, nous avons vu plus haut que les fonctions dictateurs sont extrémales. Le cas des fonctions dictateurs est un cas extrême où une coordonnée a beaucoup d'influence. Il est intéressant (surtout si l'on pense à des systèmes de votes) de se restreindre a la classe de fonctions booléennes dont toutes les influences tendent vers 0 quand n tend vers l'infini. Sous cette condition, cette question s'avère beaucoup plus intéressante. Notons que pour l'espace gaussien, le théorème de Borell répond à la question même pour des ensembles de petites influences géométriques. La différence fondamentale est que sur le cube, il n'y a plus d'invariance par rotations et les demi espaces du cubes peuvent avoir des comportements très différents.

La section suivante est consacrée à la présentation du théorème “majority is stablest”, qui comme on le verra, présente des analogies avec les problèmes vus dans le cadre continu en première partie.

4.7 Théorie du choix social et géométrie gaussienne.

Par commodité, on prendra ici $\{0,1\}^n$ pour le cube booléen, et des fonctions booléennes sont des fonctions à valeurs dans $\{0,1\}$. La fonction majorité s'écrit alors $x \mapsto \mathbf{1}_{\{\sum_{i=1}^n x_i \geq n/2\}}(x)$, où par simplicité on suppose que n est impair. Grâce au théorème de la limite centrale, on peut calculer la limite de la “noise stablilty” des fonctions Maj_n quand n tend vers l'infini. Cette valeur a été calculée dès le XIXe siècle par Sheppard [She].

4.7.1 La stabilité au bruit de la Majorité.

On rappelle que la stabilité au bruit η de la fonction majorité de taille n est par définition $\mathscr{S}_{\eta,b}(\mathrm{Maj_n}) = \mathbb{E}_\nu[\mathrm{Maj_n}(X)\mathrm{Maj_n}(Y)]$, pour des variables X et Y η-corrélées. Or on peut remplacer $\{\sum_{i=1}^n x_i \geq n/2\}$ par $\{\frac{1}{\sqrt{n}}\sum_{i=1}^n (x_i - 1/2) \geq 0\}$. Comme $\frac{1}{\sqrt{n}}\sum_{i=1}^n (x_i - 1/2)$ converge en loi vers une loi normale centrée réduite,

$$\lim_{n\to\infty} \mathscr{S}_{\eta,b}(\mathrm{Maj_n}) = \mathbb{E}_\eta[\mathbf{1}_{\{X\geq 0\}}\mathbf{1}_{\{Y\geq 0\}}],$$

où l'espérance est prise sous la loi $\mathscr{N}(0,\Gamma)$ avec $\Gamma = \begin{pmatrix} 1 & \eta \\ \eta & 1 \end{pmatrix}$.

Ainsi, par invariance par rotation de la loi $\mathscr{N}(0,\Gamma)$, cette dernière quantité vaut donc $J_\eta^B(1/2,1/2)$. On rappelle que

$$J_\eta^B(1/2,1/2) = \int_{-\infty}^0 \int_{-\infty}^0 p_\eta(x,y)d\gamma_1(x)d\gamma_1(y) = \frac{1}{2\pi\sqrt{1-\eta^2}} \int_{-\infty}^0 \int_{-\infty}^0 e^{-\frac{x^2-2\eta xy+y^2}{2(1-\eta^2)}}dxdy.$$

Or

$$\begin{aligned}
\frac{1}{2\pi\sqrt{1-\eta^2}} \int_{-\infty}^0 \int_{-\infty}^0 e^{-\frac{x^2-2\eta xy+y^2}{2(1-\eta^2)}}dxdy &= \frac{1}{\sqrt{2\pi(1-\eta^2)}} \int_{-\infty}^0 \int_{-\infty}^0 e^{-\frac{(y-\eta x)^2}{2(1-\eta^2)}}dyd\gamma_1(x) \\
&= \int_0^\infty \Phi\Big(\frac{\eta}{\sqrt{1-\eta^2}}x\Big)d\gamma_1(x).
\end{aligned}$$

Cette dernière intégrale se calcule facilement en remarquant que si on pose $\Psi : \alpha \mapsto \int_0^\infty \Phi(\alpha x)d\gamma_1(x)$, alors $\Psi'(\alpha) = \frac{1}{2\pi}\int_0^\infty xe^{-(\alpha^2+1)x^2/2}dx = \frac{1}{2\pi(\alpha^2+1)}$ d'où l'on tire que $\Psi(\alpha) = \frac{1}{4} + \frac{\arctan(\alpha)}{2\pi}$. Ainsi, si $\alpha = \frac{\eta}{\sqrt{1-\eta^2}}$,

$$J_\eta^B(1/2,1/2) = \frac{1}{4} + \frac{\arctan(\alpha)}{2\pi} = \frac{1}{4} + \frac{\arcsin(\eta)}{2\pi}.$$

Cette dernière fonction de η sera désormais appelée stabilité au bruit de la fonction majorité. En utilisant la concavité de la fonction arcsinus, on retrouve bien que cette valeur est inférieure à celle correspondant à des fonctions dictateurs (qui vaut alors avec cette définition du cube $\frac{1+\eta}{4}$). Si l'on impose que toutes les influences tendent vers 0, en revanche, le problème de maximiser la stabilité au bruit devient non trivial.

4.7.2 Le théorème "Majority is stablest"

Un problème célèbre appelé "Max-Cut", en algorithmique et en théorie des graphes, consiste à trouver le nombre maximal de coupures. Ce problème est un problème NP-complet de théorie des graphes. Il admet, sous la conjecture de jeu unique émise par Khot en 2004, une approximation par un algorithme polynomial dont l'ordre est donné par la stabilité au bruit de la fonction Majorité. Khot et al. [K-K-O-M] conjecturent que la fonction majorité maximise essentiellement la stabilité au bruit, au sens où elle est extrémale à un terme d'erreur près qui tend vers 0 quand le maximum des influences tend vers 0. L'application principale de cette conjecture est la capacité à optimiser la proportion d'approximation algorithmique du problème Max-Cut. La conjecture est prouvée peu après par Mossel, O'Donnell et Oleskiewicz [M-O-O] et porte le nom de Théorème "Majority is Stablest".

Théorème 4.7.1 (Théorème "Majority is Stablest")**.** *Soit* $f : \{0,1\}^n \to [0,1]$ *tel que* $\max_i I_i(f) \le \tau$*, alors*

$$\mathscr{S}_{\eta,b}(f) \le J_\eta^B(\mathbb{E}_\nu(f),\mathbb{E}_\nu(f)) + C(\eta)\left(\frac{\log\log 1/\tau}{\log 1/\tau}\right).$$

Rappelons que sur l'espace gaussien, on a $\mathscr{S}_\eta(A) \le \mathscr{S}_\eta(H) = J_\eta^B(\mathbb{E}_\nu(\mathbf{1}_A),\mathbb{E}_\nu(\mathbf{1}_A))$, c'est-à-dire l'inégalité du théorème sans terme d'erreur. Le théorème "Majority is Stablest" s'apparente ainsi à une version analogue sur le cube discret du résultat de Borell. Il n'est pas surprenant qu'en effet, les deux preuves connues à ce jour de ce théorème utilise le résultat de Borell. La première preuve de [M-O-O] l'utilise ainsi qu'un "principe d'invariance", une sorte de vaste généralisation du théorème de Berry-Esseen.

Rappelons que la décomposition en polynômes multilinéaires $x \in \mathbb{R}^n \mapsto \sum_S \hat{f}(S)W_S(x)$ permet de relever toute fonction f du cube discret $\{0,1\}^n$ dans $\mathbb{R}^n$. Comme $W_S(x)$ est précisément $\prod_{i\in S} h_1(x_i)$ où h_1 correspond au premier polynôme de Hermite, $Q_t h_S$ et $T_t W_S$ agissent de la même façon sur la fonction relevée, que nous noterons de la même façon f. De plus, si la fonction possède de petites influences, les distributions de $f(X)$ et de $f(G)$ sont alors proches, si $X \sim \nu$ et $G \sim \mathscr{N}(0,1)$. Plus précisément, voici un énoncé du "principe d'invariance" de Mossel, O'Donnell et Oleskiewicz.

Théorème 4.7.2. *Soit ψ une fonction C-Lipschitz. Soit f dont les influences sont plus petites que τ et de degré au plus d. Alors*

$$|\mathbb{E}_\nu(\psi(f)) - \mathbb{E}_\gamma(\psi(f))| \le O(C)2^d\tau^{1/4}. \tag{4.4}$$

Ce théorème permet de relever une fonction booléenne sur $\mathbb{R}^n$ tout en conservant ces propriétés à une petite erreur près. En utilisant le théorème de Borell pour la fonction relevée, on peut alors conclure. Montrons succinctement comment on peut atteindre par ce principe le théorème "Majority is stablest".

Posons $\eta = e^{-t}$. L'idée est d'utiliser le principe d'invariance à $\psi(T_t f) = (T_t f)^2$, qui n'est évidemment pas lipschitzienne. Mais comme f est à valeurs dans $[0,1]$, on peut définir une fonction lipschitzienne sur $\mathbb{R}$ Sq telle que $\mathbb{E}_\nu(\mathrm{Sq}(f)) = \mathscr{S}_\eta(f)$. Précisément

$$\mathrm{Sq} : x \in \mathbb{R} \mapsto x^2\mathbf{1}_{[0,1]}(x) + \mathbf{1}_{[1,\infty)}(x).$$

On a alors, par (4.4),

$$|\mathscr{S}_{\eta,b}(f) - \mathbb{E}_\gamma(\mathrm{Sq}Q_t f)| = |\mathbb{E}_\nu(\mathrm{Sq}T_t f) - \mathbb{E}_\gamma(\mathrm{Sq}Q_t f)| \le O(C)2^d\tau^{1/4}.$$

Une seconde difficulté est que la fonction relevée f sur $\mathbb{R}^n$ n'est plus forcément à valeur dans $[0,1]$ ce qui empêche d'appliquer directement le théorème de Borell. Si on définit $F = \max(0, \min(1, f))$, alors $|F - f| = \mathrm{dist}_{[0,1]}F$, et $\mathrm{dist}_{[0,1]}$ est une fonction 1 lipschitzienne. Comme $(Q_t)_{t\geq 0}$ est une contraction et Sq est 2 lipschitzienne,

$$|\mathbb{E}_\gamma(\mathrm{Sq}Q_t f) - \mathbb{E}_\gamma(\mathrm{Sq}Q_t F)| = |\mathbb{E}_\gamma(\mathrm{Sq}Q_t f) - \mathscr{S}_\eta(F)| \leq 2\mathbb{E}_\gamma(|Q_t f - Q_t F|) \leq 2\mathbb{E}_\gamma(\mathrm{dist}_{[0,1]}F).$$

De plus, comme f et F coïncident sur $[0,1]$, on a par (4.4),

$$|\mathbb{E}_\nu(\mathrm{dist}_{[0,1]}F) - \mathbb{E}_\gamma(\mathrm{dist}_{[0,1]}F)| = \mathbb{E}_\gamma(\mathrm{dist}_{[0,1]}F) \leq O(C)2^d\tau^{1/4}.$$

De plus, par l'inégalité de Borell, $\mathscr{S}_\eta(F) \leq J_\eta^B(\mathbb{E}_\gamma F, \mathbb{E}_\gamma F)$. L'inégalité ci-dessus implique que

$$|\mathbb{E}_\gamma F - \mathbb{E}_\gamma f| \leq \mathbb{E}_\gamma(\mathrm{dist}_{[0,1]}F) \leq O(C)2^d\tau^{1/4},$$

et comme $\mu \mapsto J_\eta^B(\mu, \mu)$ est $C(\eta)$-lipschitzienne, ceci prouve que

$$\mathscr{S}_{\eta,b}(f) \leq J_\eta^B(\mathbb{E}_\nu f, \mathbb{E}_\nu f) + C(\eta)2^d\tau^{1/4}.$$

Pour une fonction f générale, on considère sa restriction au d premiers degrés $f^{\leq d}$. On utilise l'inégalité précédente pour $f^{\leq d}$, puis on choisit d en fonction de τ pour optimiser les paramètres. On aboutit ainsi à la borne suivante, dont le terme d'erreur tend bien vers 0 avec τ :

$$\mathscr{S}_{\eta,b}(f) \leq J_\eta^B(\mathbb{E}_\nu f, \mathbb{E}_\nu f) + C(\eta)\left(\frac{\log\log 1/\tau}{\log 1/\tau}\right).$$

Mentionnons que dans un travail récent, Filmus, Kindler, Mossel et Wimmer sont capables de démontrer ce théorème sur le modèle de Bernouilli–Laplace en établissant sur ce graphe un principe d'invariance analogue à celui de Mossel O'Donnell et Oleskiewicz.

Quelques années plus tard, De, Mossel et Neeman [D-M-Ne] redémontrent ce résultat par un argument de tensorisation à partir d'une inégalité sur l'espace à deux points, procédant ainsi de manière analogue à Bobkov pour sa preuve sur l'isopérimétrie gaussienne. Nous donnons une preuve de De, Mossel et Neeman dans le cadre des mesures biaisées.

Rappelons que la forme fonctionnelle due à Mossel et Neeman du théorème de Borell s'écrit

$$\mathbb{E}_\eta(J_\eta^B(f(X), g(Y))) \leq J_\eta^B(\mathbb{E}_\gamma f, \mathbb{E}_\gamma g),$$

avec (X, Y) des variables gaussiennes η-corrélées. L'énoncé analogue sur le cube discret n'est pas vrai, car les fonctions dictateurs sont des contre exemples comme vu plus haut. Cependant, la conclusion du théorème “Majority is Stablest” est que l'inégalité ci-dessus est “presque vraie” sur le cube sous la condition qu'aucune influence ne soit prédominante, i.e. que toutes les influences tendent vers 0 avec n. Comme $J_\eta^B(x, y) \geq xy$, l'énoncé de “Majority is Stablest” se déduit de l'inégalité suivante :

$$\mathbb{E}_\eta(J_\eta^B(f(X), g(Y))) \leq J_\eta^B(\mathbb{E}_\nu f, \mathbb{E}_\nu g) + o_\tau(1),$$

ce qui est précisément établi dans [D-M-Ne]. La preuve repose sur une inégalité établie sur le cube discret obtenue à partir d'une inégalité sur l'espace à deux points et d'un argument de tensorisation. L'inégalité dans [D-M-Ne] est obtenue pour la mesure uniforme, mais elle s'obtient de la même manière sur le cube biaisé. Notons $\kappa_\eta = \kappa_{e^{-t}} = k_t^p$ le noyau du semi-groupe $(T_t^p)_{t\geq 0} = (e^{tL})_{t\geq 0}$ où l'on rappelle que $Lf = \int_{\{-1,1\}} f\, d\nu_p - f$. Alors $t \mapsto k_t^p$ est solution de l'équation différentielle $y' = 1 - y$, et ainsi $k_t^p(x,y) = (k_0^p(x,y) - 1)e^{-t} + 1$ avec les conditions initiales données par $f(x) = \int_\Omega f(y)k_0^p(x,y)d\nu_p(y)$, ce qui donne $k_0^p(1,1) = 1/p$, $k_0^p(-1,-1) = 1/(1-p)$ et $k_0^p(1,-1) = k_0^p(-1,1) = 0$.

Notons $f : x \in \{-1,1\} \mapsto a+ux$, $g : y \in \{-1,1\} \mapsto b+vy$, $\mathbb{E}_\eta J_\eta^B(f,g)$ s'écrit alors explicitement comme

$$\begin{aligned}&(p^2+p(1-p)\eta)J_\eta^B(a+u,b+v)+((1-p)^2+p(1-p)\eta)J_\eta^B(a-u,b-v)+\\&p(1-p)(1-\eta)(J_\eta^B(a+u,b-v)+J_\eta^B(a-u,b+v)).\end{aligned}$$

De plus,

$$J_\eta^B(\mathbb{E}f,\mathbb{E}g)=J_\eta^B(a+(2p-1)u,\ b+(2p-1)v).$$

Par un développement de Taylor à l'ordre 3 en (u,v), on obtient :

$$\mathbb{E}_\eta J_\eta^B(f,g)-J_\eta^B(\mathbb{E}f,\mathbb{E}g)=A_p(\partial_{11}J_\eta^B(a,b)u^2+2\eta\partial_{12}J_\eta^B(a,b)uv+\partial_{22}J_\eta^B(a,b)v^2)+O(|u|^3,|v|^3,\eta)$$

où $A_p=2p(1-p)$. On rappelle que J_η^B est η-concave, c'est-à-dire

$$M_\eta=\begin{pmatrix}\partial_{11}J_\eta^B & \eta\partial_{12}J_\eta^B\\ \eta\partial_{12}J_\eta^B & \partial_{22}J_\eta^B\end{pmatrix}\leq 0.$$

De plus, quelques calculs montrent qu'il existe des constantes $C(\eta)$ et $c(\eta)$ telles que :

$$\forall\, i+j=3,\ |\partial_{i,j}^3 J_\eta^B(x,y)|\leq C(\eta)\left(\frac{1}{x(1-x)y(1-y)}\right)^{c(\eta)}.$$

Ainsi sur l'espace à deux points, on a "l'inégalité de Borell discrète" pour des fonctions à valeurs dans $[\varepsilon,1-\varepsilon]$:

$$\mathbb{E}_\eta J_\eta^B(f,g)\leq J_\eta^B\left(\mathbb{E}_{\nu_p}f,\mathbb{E}_{\nu_p}g\right)+C(\eta,p)\varepsilon^{-c(\eta)}(|f-\mathbb{E}_{\nu_p}f|^3+|g-\mathbb{E}_{\nu_p}g|^3). \tag{4.5}$$

Cette équation se tensorise sous la forme suivante :

$$\mathbb{E}_\eta J_\eta^B(f,g)\leq J_\eta^B\left(\mathbb{E}_{\nu_p^{\otimes n}}f,\mathbb{E}_{\nu_p^{\otimes n}}g\right)+C(\eta,p)\varepsilon^{-c(\eta)}(\Delta_n(f)+\Delta_n(g)), \tag{4.6}$$

avec $\Delta_n(f)$ défini par récurrence comme $\Delta_1(f)=\mathbb{E}_{\nu_p}|f-\mathbb{E}_{\nu_p}f|^3$ et $\Delta_n(f)=\mathbb{E}_{X_n}(\Delta_{n-1}f_{X_n})+\Delta_1(\mathbb{E}(f_{X_n}|X_n)$. Pour $f=g$, puisque $J_\eta^B(x,y)\geq xy$, (4.6) implique que :

$$\mathscr{S}_\eta^p(f)\leq J_\eta^B\left(\mathbb{E}_{\nu_p^{\otimes n}}f,\mathbb{E}_{\nu_p^{\otimes n}}g\right)+C(\eta,p)\varepsilon^{-c(\eta)}\Delta_n(f).$$

La tâche principale de [D-M-Ne] est donc de borner le terme d'erreur $C(\eta,p)\varepsilon^{-c(\eta)}\Delta_n(f)$ par le maximum des influences de f. Or dans le cas biaisé,

$$\mathbb{E}(|f-\mathbb{E}_{\nu_p}(f)|^3)=\frac{p^2+(1-p)^2}{\sqrt{p(1-p)}}|\hat{f}(1)|^3=K_p|\hat{f}(1)|^3.$$

Par récurrence, en notant $S_i=\{i+1,\ldots,n\}$, on obtient

$$\Delta_n(f)=K_p\sum_{i=1}^n\mathbb{E}_{X\in\{-1,1\}^{S_i}}|\hat{f}_X(i)|^3=K_p\sum_{S\in[n],S\neq\emptyset}\hat{f}^3(S).$$

En réécrivant

$$\sum_{S\in[n],S\neq\emptyset}\hat{f}^3(S)=\sum_{i=1}^n\mathbb{E}_{X\in\{-1,1\}^{S_i}}|\hat{f}_X(i)|^3$$

l'idée est d'utiliser l'inégalité (4.6) à une petite perturbation de f, pour pouvoir utiliser l'hypercontractivité. Par abus de notation, on note $T^p_{1-\eta_1}f$ pour $T^p_{-\log(1-\eta_1)}f$. Pour η_1 petit, cette perturbation n'affecte pas trop la stabilité au bruit car

$$|\mathscr{S}^p_\eta(f)-\mathscr{S}^p_\eta(T^p_{1-\eta_1}f)| = |\mathscr{S}^p_\eta(f)-\mathscr{S}^p_{(1-\eta_1)^2\eta}(f)| \leq O(\eta_1). \tag{4.7}$$

De plus, l'hypercontracitivité du semi-groupe (T^p_t) implique que

$$\sum_{i=1}^n \mathbb{E}_{X\in\{-1,1\}^{S_i}} |\widehat{T^p_{1-\eta_1}f}_X(i)|^r \leq \left(\sum_{i=1}^n \mathbb{E}_{X\in\{-1,1\}^{S_i}} |\hat{f}_X(i)|^2\right)^{r/2} \tag{4.8}$$

avec $r = 1+(1-\eta_1)^{-2C_p}$, où l'on rappelle que $C_p = 2\frac{2p-1}{\log p - \log(1-p)}$. Comme pour tout $i \in \{1,\dots,n\}$,

$$\mathbb{E}_{X\in\{-1,1\}^{S_i}} |\hat{f}_X(i)|^2 \leq \sum_{i\in S} \hat{f}^2(S) = I_i(f) \leq \tau,$$

et comme l'idendité de Parseval implique que

$$\sum_{i=1}^n \mathbb{E}_{X\in\{-1,1\}^{S_i}} |\hat{f}_X(i)|^2 = \sum_{S\in[n],S\neq\emptyset} \hat{f}^2(S) = \mathrm{Var}_{\nu_p}(f) \leq 1,$$

on a par (4.8) la majoration suivante :

$$\sum_{i=1}^n \mathbb{E}_{X\in\{-1,1\}^{S_i}} |\widehat{T^p_{1-\eta_1}f}_X(i)|^r \leq \left(\sum_{i=1}^n \mathbb{E}_{X\in\{-1,1\}^{S_i}} |\widehat{T^p_{1-\eta_1}f}_X(i)|^2\right)\tau^{\frac{r-2}{2}} \leq \tau^{\frac{r-2}{2}}.$$

Ainsi, si η_1 est assez petit, cela implique que $r \leq 3$ et, de plus, il existe une constante numérique $c \in (0,1)$ telle que :

$$\Delta_n(T^p_{1-\eta_1}f) \leq K_p \sum_{i=1}^n \mathbb{E}_{X\in\{-1,1\}^{S_i}} |\widehat{(T^p_{1-\eta_1}f)}_X(i)|^r \leq K_p \tau^{cC_p\eta_1},$$

et alors

$$\mathscr{S}^p_{(1-\eta_1)^2\eta}(f) \leq J^B_\eta(\mathbb{E}(f),\mathbb{E}(f)) + C(\eta,p)\varepsilon^{-c(\eta)}\tau^{cC_p\eta_1}. \tag{4.9}$$

Lorsque p est indépendant de n, on peut enlever les dépendances en p.

Soit maintenant une fonction générale $f : \{0,1\}^n \to [0,1]$. On tronque la fonction sur $[\varepsilon, 1-\varepsilon]$ en prenant $f^+_\varepsilon = \max(\varepsilon, f)$ et ensuite $f_\varepsilon = \min(1-\varepsilon, f^+_\varepsilon)$. On lui applique une petite perturbation $T^p_{1-\eta_1}$, et par les inégalités (4.7), (4.10) et le fait que $\mu \mapsto J^B_\eta(\mu,\mu)$ est $C(\eta)$-lipschitzienne,

$$\mathscr{S}^p_\eta(f) \leq \mathscr{S}^p_{(1-\eta_1)^2\eta}(f_\varepsilon) + C(\varepsilon+\eta_1) \leq J^B_\eta(\mathbb{E}(f),\mathbb{E}(f)) + C(\eta)(\varepsilon+\eta_1+\varepsilon^{-c(\eta)}\tau^{c\eta_1}).$$

En choisissant η_1 et ε de manière appropriée, on en déduit ainsi

$$\mathscr{S}^p_\eta(f) \leq J^B_\eta(\mathbb{E}(f),\mathbb{E}(f)) + C(\eta)\left(\frac{\log\log 1/\tau}{\log 1/\tau}\right). \tag{4.10}$$

Ceci est la version biaisée du théorème “Majority is stablest”.

4.7.3 Remarque sur les espaces discrets généraux et résultat sur les tranches du cube

Une question en grande partie ouverte est celle de la généralisation de ce théorème sur des modèles discrets plus généraux. Nous avons en effet défini les notions de stabilité au bruit et d'influences sur de tels espaces. La question plutôt générale est de savoir pour quels graphes (G_n, V_n), munis de leurs mesures uniformes μ_n, l'inégalité suivante est vraie pour tout $f : G_n \to [0,1]$ telles que $\max_{s\in S_n} I_s(f) = \tau$,

$$\mathcal{S}_\eta(f) \leq J_\eta^B(\mathbb{E}_{\mu_n}(f), \mathbb{E}_{\mu_n}(f)) + o_\tau(1).$$

À ce jour, il existe une généralisation autre que sur le cube de ce résultat, sur les tranches du cube booléen d'ordre k $\binom{[n]}{k} := \{x \in \{0,1\}^n, \sum_{i=1}^n x_i = k\}$, avec k de l'ordre de n. Rappelons que sur cet espace, l'inégalité de Talagrand est vraie. Notons $\mathrm{S}_\eta^s(f)$ la “noise stability” dans ce contexte. On rappelle que les influences sont définies par

$$I_{ij}f := \|x \mapsto f(x) - f(x^{(ij)})\|_{L^1(\mu)} = \|D_{(ij)}f\|_{L^1(\mu)}$$

où μ est la mesure uniforme sur $\binom{[n]}{k}$. Alors on a le théorème suivant :

Théorème 4.7.3. *Pour toute fonction définie sur $\binom{[n]}{k}$ à valeurs dans $[0,1]$ et dont les influences sont bornées par τ, on a*

$$\mathrm{S}_\eta^s(f) \leq J_\eta^B(\mathbb{E}_\mu(f), \mathbb{E}_\mu(f)) + o_\tau(1).$$

Ce théorème est en fait un corollaire de l'extension du principe d'invariance aux tranches du cube. Une fois ce principe établi, la preuve suit par les mêmes arguments, rappelés plus haut, que ceux de [M-O-O].

Ce principe d'invariance s'appuie sur le travail préalable de Filmus [Fil]. Dans [Fil], Filmus construit une base orthonormée de polynômes multilinéaires pour les tranches du cube. Ceci permet de relever toute fonction du cube sur $\mathbb{R}^n$ et donc également sur le cube booléen. Notons pour f définie sur les tranches du cube, $\tilde{f}$ la fonction relevée sur le cube. Cette opération préserve l'espérance, c'est-à-dire que si $p = k/n$, $\mathbb{E}_\mu f = \mathbb{E}_{\nu_p}\tilde{f}$. De plus comme dans le cas du cube, la base $(\chi_a)_a$ est telle que les éléments qui la constituent sont des fonctions propres du semi-groupe sous-jacent $(P_t)_{t\geq 0}$. Ainsi, de manière spectrale, on peut écrire

$$\mathcal{S}_\eta^s(f) = \sum_{d=0}^{\deg f} \eta^{c_k}\|f^{=k}\|_2^2,$$

où cette fois $c_k = k(1-(k-1)/n)$. Contrairement aux cas du cube et de l'espace gaussien, les semi-groupes n'agissent pas de la même façon sur χ_a. Toutefois, si n est assez grand, c'est “presque” le cas. En effet,

$$\mathcal{S}_\eta^s(f) - \mathcal{S}_\eta^p(\tilde{f}) = \sum_{k=0}^{\deg f}(\eta^{c_k} - \eta^k)\|f^{=k}\|_2^2,$$

et comme pour tout $k \leq n/2$, $\eta > 1/2$, $\eta^{c_k} - \eta^k \leq \frac{C(\eta)}{n}$, on a pour toute fonction définie sur les tranches du cube à valeurs dans $[0,1]$

$$\mathcal{S}_\eta^s(f) - \mathcal{S}_\eta^p(\tilde{f}) \leq \frac{C(\eta)}{n}.$$

De même, les influences de f et $\tilde{f}$ sont presque du même ordre si le degré de f est bas. Nous avons en effet le lemme suivant :

Lemme 4.7.4. *Pour toute fonction f de degré au plus d définie sur $\binom{[n]}{k}$, si l'on pose $I_i(f) = \frac{1}{n}\sum_{j\neq i} I_{ij}(f)$, alors*

$$I_i(\tilde{f}) \leq \frac{2d}{p(1-p)(n-d)}\mathrm{Var}_\mu(f) + \frac{2n}{p(1-p)(n-d)}I_i(f).$$

La contrainte de se restreindre à des fonctions de faible degré n'est pas trop restrictive. En effet, on peut prendre $d = n^{1/3}$ plus haut. La restriction d'une fonction générale à ses termes de degré au plus $n^{1/3}$ n'affecte sa noise stability et son espérance que d'un terme négligeable $o_n(1)$. La difficulté en vue d'appliquer (4.10) est que l'extension $\tilde{f}$ n'est pas forcément à valeur dans $[0,1]$.

Si l'on définit $F = \max(0, \min(1, \tilde{f}))$, il reste à démontrer que les quantités $\mathscr{S}_\eta^p(\tilde{f}) - \mathscr{S}_{\eta,p}(F)$ et $\mathbb{E}_{\nu_p} F - \mathbb{E}_{\nu_p}(\tilde{f})$ sont petites. Le principe d'invariance le garantit, mais il est possible que l'on puisse obtenir ce résultat de manière plus directe.

Pour en revenir à la généralisation possible sur des modèles plus généraux, deux difficultés se présentent. La preuve discrète de De Mossel et Neeman, bien que basée sur une idée simple, nécessite d'écrire la première ligne. Même sur les tranches du cube booléen, ce n'est pas clair.

La preuve par principe d'invariance, quant à elle, nécessite la bonne connaissance d'une base orthonormée sur laquelle décomposer les fonctions selon ses "degrés".

4.7.4 Remarque sur l'hypercontractivité.

Pour conclure cette section, faisons une remarque sur l'hypercontractivité. Comme expliqué dans le cas gaussien, un autre exemple de fonction η-concave est donné par fonction d'hypercontractivité $J_\eta^H(u,v) = u^\alpha v^\beta$ avec $\alpha\beta\eta^2 = (\alpha-1)(\beta-1)$. Le lemme de Bonami sur deux points $\|T_t f\|_{L^p(\{-1,1\})} \leq \|f\|_{L^q(\{-1,1)\})}$ implique par dualité :

$$\mathbb{E}_\eta J_\eta^H(f,g) \leq J_\eta^H(\mathbb{E}_\nu f, \mathbb{E}_\nu g),$$

et ainsi par tensorisation, l'hypercontractivité dualisée s'écrit :

$$\mathbb{E}_\eta J_\eta^H(f,g) \leq J_\eta^H(\mathbb{E}_{\nu^{\otimes n}} f, \mathbb{E}_{\nu^{\otimes n}} g).$$

Cela signifie que contrairement à la fonction de Borell J_η^B, l'inégalité pour J_η^H s'écrit de la même façon que pour le cas gaussien. Cette propriété est ainsi strictement plus forte que celle d'η-concavité. En fait, le cas à deux points s'écrit :

$$(1+\eta)[(1+u)^\alpha(1+v)^\beta + (1-u)^\alpha(1-v)^\beta] + (1-\eta)[(1-u)^\alpha(1+v)^\beta + (1+u)^\alpha(1-v)^\beta] \leq 4.$$

Par homogénéité, u, v peuvent être supposés dans $(0,1)$. Contrairement au cas non dualisé, cette inégalité s'avère de manière surprenante beaucoup plus délicate à prouver. En développant en série entière, elle se réécrit :

$$\sum_{k\geq 1} \frac{2}{(2k)!}\left(a_k v^{2(k+1)} + b_k \eta v^{2k}\right) \leq 0,$$

avec

$$a_k = \sum_{0\leq j<k} \binom{2k}{2j+1}\binom{\alpha}{2j+1}\binom{\beta}{2(k-j)+1}\left(\frac{u}{v}\right)^{2j+1} \qquad b_k = \sum_{0\leq m\leq k} \binom{2k}{2m}\binom{\alpha}{2m}\binom{\beta}{2(k-m)}\left(\frac{u}{v}\right)^{2m}$$

et contrairement au cas à deux points, tous les termes du développements ne sont pas nécessairement négatifs.

Notons que nous avons trouvé une preuve directe de cette inégalité par Schramm and Tsirelson [Sc-Ts] (écrite dans le cas où $\alpha = \beta$). Ce lemme sur deux points est utilisé par les auteurs pour donner une preuve à l'inégalité d'hypercontractivité sur des modèles d'arbres en utilisant un argument de martingale.

4.8 Présentation des travaux de la seconde partie.

Cette section a pour but de présenter les travaux réalisés dans l'article [Bou]. Le résultat principal est l'extension d'une version quantitative du critère de Benjamini, Kalai et Schramm à des modèles de graphes de Schreier assez généraux, ainsi que sa version sur des modèles continus. L'analogue continu nécessite la définition d'influences et de stabilité au bruit dans ce cadre. Nous nous appuyons sur la définition d'influences géométriques de Keller, Mossel et Sen, issue de [K-M-S1], qui est présentée dans la sous-section suivante.

Avant d'aborder le critère de Benjamini, Kalai et Schramm, il nous semble bon de présenter succinctement les extensions faites par O'Donnell et Wimmer et Cordero-Erausquin et Ledoux des inégalités de Kahn, Kalai et Linial et de Talagrand sur des modèles de graphes de Cayley ou de Schreier. En effet, le schéma de preuve de nos résultats est directement inspiré de celui de [CE-L]. L'article [O-W] présente quant à lui plusieurs intérêts. En premier lieu, les auteurs y donnent un critère sur les graphes pour avoir une commutation entre le semi-groupe et les influences. Cette propriété est cruciale et est présente dans tous les travaux cités. De plus, cette preuve fait apparaître un lien entre la constante de Sobolev logarithmique et celle de trou spectral pour qu'une inégalité de ce type puisse améliorer l'inégalité de Poincaré. Enfin, les idées de la preuve d'O-Donnell et Wimmer ont été reprises par Förstromm qui de manière indépendante à notre travail a également étendu le critère B-K-S sur des graphes de Schreier.

Influences géométriques

Soit $(\mathbb{R}^n, \mu^{\otimes n})$ un espace probabilisé. Dans ce cadre continu, Keller, Mossel et Sen [K-M-S1] définissent la notion d'influence géométrique d'une coordonnée i, $i \in \{1, \ldots, n\}$ par

$$I_i^{\mathcal{G}}(A) = \mathbb{E}_x[\mu^+(A_i^x)].$$

Ici $A_i^x \subset \mathbb{R}$ est la restriction de A le long de la fibre $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, c'est-à-dire

$$A_i^x = \{y \in \mathbb{R}, (x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n) \in A\}$$

et où l'on rappelle que μ^+ dénote le contenu inférieur de Minkowski. De manière heuristique, $I_i^{\mathcal{G}}(A) = \mathbb{E}_\mu(|\partial_i \mathbf{1}_A|)$. Plus précisément, comme démontré par les auteurs, lorsque l'ensemble A est monotone (i.e. $\mathbf{1}_A$ est monotone), $I_i^{\mathcal{G}}(A) = \lim_{f \to \mathbf{1}_A} \mathbb{E}_\mu(|\partial_i f|)$. Concernant cette définition d'influences géométriques, plusieurs remarques peuvent être faites. Pour un sous-ensemble $A \subset \mathbb{R}^n$ croissant ou décroissant, la somme $\sum_{i=1}^n I_i^{\mathcal{G}}(A)$ représente la mesure de bord pour l'élargissement uniforme, c'est-à-dire

$$\mu_\infty^+(A) = \liminf_{r \to 0} \frac{\mu(A + [-r, r]^n) - \mu(A)}{r}.$$

Lorsque A est un sous-ensemble dont le bord ∂A est une surface régulière dont nous notons σ_A sa mesure de surface, si ρ désigne la densité de μ et n_A la normale extérieure à A, on a :

$$\mu_\infty^+(A) = \int_{\partial A} \|n_A(x)\|_1 \rho(x) d\sigma_A(x).$$

Remarquons que comme $\|n_A\|_1 \geq 1$, on a toujours $\mu_\infty^+(A) \geq \mu^+(A)$.

Les auteurs prouvent un analogue du théorème de Kahn, Kalai et Linial pour des mesures de Boltzmann de la forme $\mu(dx) = \frac{1}{Z_\beta} e^{-|x|^\beta} dx$, c'est-à-dire $\max_i I_i^{\mathcal{G}}(A) \geq \frac{\log n}{n}$. La preuve repose essentiellement sur des arguments géométriques.

Rappelons que l'inégalité de Talagrand implique celle de Kahn, Kalai et Linial. L'extension de l'inégalité de Talagrand sur l'espace gaussien - et plus généralement sur des mesures log-concaves, bien que non publiée, était connu depuis longtemps par les spécialistes. Un obstacle à l'application de cette inégalité aux influences géométriques est que $\mathbb{E}_\mu(\partial_i f)^2$ ne converge pas vers la i-ème influence géométrique d'un ensemble lorsque f approche la fonction caractéristique $\mathbf{1}_A$. Dans [CE-L], les auteurs modifient donc l'inégalité de Talagrand de façon à n'exprimer que des normes L^1 de dérivées. L'inégalité qui en découle permet de retrouver l'analogue du théorème KKL pour les mesures $\mu(dx) = \frac{1}{Z_\beta} e^{-|x|^\beta} dx$, mais seulement pour $\beta \geq 2$.

Notons également qu'un travail indépendant de Barthe et Huou [B-H] sur le problème isopérimétrique sous l'élargissement uniforme permet de retrouver le théorème K-K-L de Keller, Mossel et Sen sous une forme plus intrinsèque.

Dans [K-M-S2], les auteurs lient cette notion d'influence géométrique à celle de noise stability/sensitivity en utilisant une version quantitative du critère B-K-S due à Keller et Kindler sur le cube et un argument utilisant le théorème de la limite centrale. Le résultat qui en découle est une version quantitative analogue sur l'espace gaussien. Dans nos travaux, nous donnons une preuve dont le schéma est directement inspiré de celui de [CE-L] sur l'inégalité de Talagrand. Ce schéma permet d'étendre les résultats de [K-M-S2] à des produits de mesures log-concaves satisfaisant le critère de courbure $CD(\kappa,\infty)$ de Bakry–Emery, en plus de couvrir des espaces discrets plus généraux.

Noise stability/sensitivity

De manière duale au concept de noise stability, on définit celui de noise sensitivity gaussienne par $\mathcal{NS}_\eta^{\mathcal{G}}(A) = \mathcal{S}_\eta(A) - \gamma(A)^2$. Notons que celle définition n'est pas universelle. On peut également définir la noise sensitivity par $\gamma(A) - \mathcal{S}_\eta(A)$, ce qui correspond en notation probabiliste à $\mathbb{P}(X \in A, Y \notin A)$, pour (X,Y) des variables gaussiennes η-corrélées. Nous conserverons la première définition.

Une suite d'ensembles $(A_n \subset \mathbb{R}^n)_{n\geq 0}$ est dite asymptotiquement sensible au bruit si $\mathcal{NS}_\eta^{\mathcal{G}}(A_n) \to 0$ quand $n \to \infty$. Comme pour les modèles discrets, la noise sensitivity se définit à l'aide du semi-groupe d'Ornstein–Uhlenbeck par

$$\mathcal{NS}_\eta^{\mathcal{G}}(A) = \mathrm{Var}_\gamma(Q_{t/2}\mathbf{1}_A),$$

pour $\eta = e^{-t}$. De manière plus générale, pour des mesures log-concaves, on pose donc naturellement la définition suivante :

Définition 4.8.1. *Soit $(\mathbb{R}^n, \mu)$ satisfaisant la condition $CD(\kappa,\infty)$. Alors on définit la noise sensitivity d'une fonction f par*

$$\mathcal{NS}_\eta(f) = \mathrm{Var}_\gamma(P_{t/2} f),$$

où $\eta = e^{-\kappa t}$.

Notons que si H est un demi-espace tel que $\mu(A) = \gamma(H)$, le théorème de comparaison obtenu dans le chapitre 2 implique donc que

$$\mathcal{NS}_\eta(A) \leq \mathcal{NS}_\eta^{\mathcal{G}}(H).$$

O'Donnell, Wimmer et K-K-L

Dans cette sous-section, nous donnons un apercu général de la preuve du théorème K-K-L généralisé par O'Donnel et Wimmer dans [O-W]. Soit (X,G) un graphe de Schreier. Notons $\{0 = \lambda_0, \lambda_1, \ldots, \lambda_i, , \ldots\}$, le spectre de $-L$ (avec $\lambda_1 = \lambda$ le trou spectral), et f_i les fonctions propres associées aux valeurs propres λ_i. On note également ρ la constante d'hypercontractivité du graphe et on rappelle que $\rho \leq \lambda$. De manière cruciale, les auteurs s'appuient sur une commutation entre le semi-groupe P_t et les opérateurs de dérivation D_s. Leur hypothèse est que le groupe générateur S est une réunion de classe de conjugaison. Alors, dans ce cas, on dispose du lemme suivant :

Lemme 4.8.2. *Supposons que S soit une réunion de classes de conjugaison. Alors pour tout $s \in S$, $D_s P_t = P_t D_s$.*

Démonstration. Nous allons prouver que les opérateurs K et $I - D_s$ commutent, ce qui implique la même chose pour L et D_s. Puisque $(I - D_s)f(x) = f(x^s)$, on a

$$K(I - D_s)f(x) = \mathbb{E}_{s'}[(I - D_s)f(x^{s'})] = \mathbb{E}_{s'}[f(x^{s's})] = \mathbb{E}_{s'}[f(x^{ss^{-1}s's})].$$

Comme $s' \mapsto s^{-1}s's$ est une injection sur S, et puisque S est un ensemble fini, c'est aussi une bijection. Ainsi

$$\mathbb{E}_{s'}[f(x^{ss^{-1}s's})] = \mathbb{E}_t[f(x^{st})] = Kf(x^s) = (I - D_s)(Kf)(x).$$

Comme L et D_s commutent, c'est également le cas pour $(P_t)_{t\geq 0}$ et D_s.

□

De plus, les auteurs supposent que les fonctions sont booléennes. Ce dernier point implique que pour tout $r \geq 1$, $I_s(f) = \|D_s f\|_r^r$. Ceci, combiné à l'hypercontractivité, entraîne le fait suivant

$$I_s(P_t f) = \|D_s P_t f\|_2^2 = \|P_t D_s f\|_2^2 \leq \|D_s f\|_{1+e^{-\rho t}}^2 = I_s(f)^{\frac{2}{1+e^{-2\rho t}}}.$$

On peut ainsi ainsi majorer $\sum_{s\in S} I_s(P_t f)$ en termes de $\sum_{s\in S} I_s(f)$ et du maximum des influences (grâce au facteur $2(1 + e^{-2\rho t})^{-1} > 1$). Par décomposition spectrale selon l'opérateur L, on exprime de plus la somme des influences comme

$$\sum_{s\in S} I_s(f) = \mathscr{E}(f) = \sum_{i\geq 0} \lambda_i \|f_i\|_2^2.$$

Appliquant cette égalité à $P_t f$, pour tout $\Lambda \geq 0$, on peut également minorer la somme des influences de $P_t f$ par $(\min_{i\leq\Lambda} \lambda_i e^{-t\lambda_i}) \sum_{0\leq i\leq\Lambda} \|f_i\|_2^2$.

En choisissant $t = \frac{\ln 3}{2}\frac{1}{\rho}$ et $\Lambda = \frac{2\mathscr{E}(f)}{\mathrm{Var}_\mu(f)}$, $\min_{i\leq\Lambda} \lambda_i e^{-t\lambda_i} = \Lambda 3^{-\frac{\Lambda}{\rho}}$, $\sum_{0\leq i\leq\Lambda} \|f_i\|_2^2 \geq \frac{1}{2}\mathrm{Var}_\mu(f)$, et ainsi

$$\sum_{s\in S} I_s(P_t f) \geq \mathscr{E}(f) 3^{-\frac{\Lambda}{\rho}}.$$

Mais d'autre part, $\frac{2}{1+e^{-2\rho t}} = 4/3$ donc l'hypercontractivité implique que

$$\sum_{s\in S} I_s(P_t f) \leq \sum_{s\in S} I_s(f)^{3/2} \leq \mathscr{E}(f) \max_{s\in S} \sqrt{I_s(f)}.$$

En particulier, en combinant les deux bornes, on obtient $\max_{s\in S} \sqrt{I_s(f)} \geq 3^{-\frac{\Lambda}{\rho}}$. En remplaçant Λ par son expression, cette inégalité implique qu'il existe une constante $c > 0$ telle que

$$\max_{s\in S} I_s(f) \geq c\rho \log\left(\frac{1}{\rho}\right) \mathrm{Var}_\mu(f).$$

Ce résultat est exprimé en termes de normalisation de graphes de Cayley. Il contient bien le théorème K-K-L sur le cube biaisé. Il est intéressant et important de noter que l'inégalité de trou spectral implique que $\max_{s\in S} I_s(f) \geq c\lambda \mathrm{Var}_\mu(f)$, ce qui signifie que la proposition améliore l'inégalité de trou spectral si et seulement si le ratio $\frac{\rho_n}{\lambda_n} \log\left(\frac{1}{\rho_n}\right)$ tend vers l'infini avec n.

Ce résultat n'est pas en soi surprenant : si l'hypercontractivité n'est pas assez forte, l'inégalité n'améliore pas celle de Poincaré. Rappelons que de nombreux exemples concrets sont disponibles où l'on connaît un ordre de grandeur de ces constantes. Par exemple, pour le groupe symétrique $\mathfrak{S}_n$, O'Donnell et Wimmer montrent qu'une minoration plus forte que celle donnée par le trou spectral ne peut pas être vrai. En effet, on a le lemme suivant :

Lemme 4.8.3. *Soit f_n la fonction caractéristique d'un dérangement de $\mathfrak{S}_n$, c'est-à-dire une permutation σ telle que pour tout $i \in \{1,\dots,n\}$, $\sigma(i) \neq i$. Alors $\forall n \geq 1$, $\mathrm{Var}_\mu(f_n) \geq 1/9$ et $\forall \tau_{ij} \in \mathcal{T}_n$, $I_{\tau_{ij}}(f_n) \leq \frac{2}{n}$.*

Démonstration. Par un principe d'exclusion/inclusion classique, le nombre total de dérangements est égal à $n! \sum_{k=0}^n \frac{(-1)^k}{k!}$. Ainsi $\mathbb{P}(f_n = 1) = \sum_{k=0}^n \frac{(-1)^k}{k!} \in (1/2, 1/3)$, et donc $\inf_{n\in\mathbb{N}} \mathrm{Var}_\mu(f_n) > 0$. De plus, toutes les transpositions ont mêmes influences. Or $(f(\sigma \circ (12)) \neq f(\sigma))$ si et seulement si $\sigma(1) = 2$ ou $\sigma(2) = 1$. Ceci arrive avec probabilité au plus $2/n$, et ainsi pour toute transposition $\tau_{ij} \in \mathcal{T}_n$, $I_{\tau_{ij}}(f) \leq 2/n$. □

Une remarque concernant cet article, dont l'esprit de la preuve est similaire à l'article original de Kahn, Kalai et Linial, est que les résultats sont énoncés pour des fonctions à valeurs dans $\{0,1\}$. Il est en effet crucial d'utiliser l'égalité pour tout $r \geq 1$, $I_s(f) = \|D_s f\|_r^r$, faute de quoi la preuve ne fonctionne plus. L'inégalité de Talagrand généralisée par Cordero-Erausquin et Ledoux permet de s'affranchir de cette condition. Nous présentons le schéma de leur preuve dans la sous-section suivante.

L'inégalité de Talagrand d'après Cordero-Erausquin et Ledoux

La preuve d'O-Donnell et Wimmer présentée utilise la décomposition de la forme de Dirichlet le long de "directions" données par le générateur S du graphe de Schreier ou Cayley. L'inégalité de Talagrand sous cette condition de décomposition s'adapte très bien à la condition que, en un certain sens, les directions commutent avec l'opérateur de dérivations. Ceci peut être exprimé de façon abstraite, qui s'applique immédiatement à des exemples concrets. c'est-à-dire, partant de la décomposition suivante :

$$\mathcal{E}(f,f) = \sum_{i=1}^m \int_\Omega \Gamma_i(f)^2 \, d\mu, \tag{4.11}$$

supposons que les opérateurs Γ_i commutent avec le semi-groupe au sens où il existe κ tel que $i = 1,\dots,m$

$$\Gamma_i(P_t f) \leq e^{\kappa t} P_t(\Gamma_i f). \tag{4.12}$$

Cette propriété, caractéristique du critère $\Gamma_2 \geq \kappa\Gamma$ dans le cadre continu, est néanmoins un peu différente dans le cas discret. En effet, on demande seulement une commutation selon toutes les directions et non une commutation globale. Dans les exemples de graphes de Schreier ou Cayley que nous considérons, la commutation entre l'opérateur de Markov L et les dérivées D_s découlent de l'hypothèse $S = S^{-1}$ (comme vu plus haut).

Notons toutefois qu'à partir du laplacien discret (somme aux plus proches voisins), Kozma, Klartag, Ralli et Tetali [K-K-R-T] donnent une définition du "carré du champ" et de ses itérés, et définissent ainsi une notion de courbure minorée similaire à celle de Bakry-Emery. Cette définition implique de manière analogue au cadre continu une commutation entre le semi-groupe et l'opérateur de dérivation. On pourrait ainsi considérer des produits de tels espaces discrets pour établir une généralisation de l'inégalité de Talagrand.

Pour les modèles continus, c'est le critère de Bakry-Emery qui permet cette commutation. Rappelons que pour le semi-groupe d'Ornstein-Uhlenbeck sur $\mathbb{R}^n$, $\partial_i Q_t f = e^{-t} Q_t(\partial_i f)$ c'est-à-dire (4.12) avec $\kappa = 1$. Pour les mesures log-concaves $d\mu(x) = e^{-V(x)}dx$ sur $\mathbb{R}^n$, dès que V satisfait $\mathrm{Hess}(V) \geq c$ avec $c \in \mathbb{R}^+$, pour toute fonction lisse $f : \mathbb{R}^n \to \mathbb{R}$ et tout $t \geq 0$,

$$|\nabla P_t f| \leq e^{-ct} P_t(|\nabla f|). \tag{4.13}$$

Si l'on prend un produit de telles mesures $d\mu_i(x) = e^{-V_i(x)}dx$, $i = 1,\dots,m$, tel que chaque V_i satisfasse $\mathrm{Hess}(V_i) \geq c_i$ avec $c_i \in \mathbb{R}^+$, on obtient une décomposition du type (4.11) avec $\Gamma_i = \nabla_i$ et $\kappa = \min c_i$. Notons que l'on pourrait considérer de manière identique des produits de variétés Riemanniennes à poids $(M_i, g_i, e^{-\psi_i} d\mathrm{vol})$ tels que $Ric_{g_i} + \mathrm{Hess}\psi_i \geq c_i$.

Pour l'inégalité de Talagrand, le point de départ est d'utiliser la décomposition de la variance le long de la forme de Dirichlet, c'est-à-dire

$$\mathrm{Var}_\mu(f) = 2\int_0^\infty \mathcal{E}(P_u f, P_u f)du = 2\sum_{s\in S}\int_0^\infty \|D_s P_u f\|_2^2 du.$$

Notons que sur les modèles discrets, la commutation exacte $D_s P_u$ ne permet pas de pouvoir intégrer jusqu'à l'infini. Cependant l'inégalité de trou spectral avec constante λ a pour forme équivalente :

$$\forall T>0,\ \mathrm{Var}_\mu(f) \leq \frac{1}{1-e^{-\lambda T}}\Big(\|f\|_2^2 - \|P_T f\|_2^2\Big).$$

Ceci permet de se restreindre à un domaine fini d'intégration $[0, T]$, car on a de même

$$\|f\|_2^2 - \|P_T f\|_2^2 = 2\int_0^T \sum_{s\in S}\|D_s P_u f\|_2^2 du.$$

L'utilisation de la propriété de commutation et ainsi que l'hypercontractivité du semi-groupe $(P_t)_{t\geq 0}$ implique alors (cf [CE-L] pour les détails) que :

$$\mathrm{Var}_\mu(f) \leq \frac{C}{\rho|S|}\sum_{s\in S}\frac{I_s(f)}{1+\log(I_s(f))}.$$

Nous retrouvons bien la condition $\rho\log(1/\rho) >> \lambda$ pour que cette inégalité améliore celle de Poincaré. Notons que ces arguments par semi-groupe permettent de considérer des fonctions à valeurs réelles, et ainsi la condition d'être booléenne n'est pas nécéssaire.

L'argument n'est pas exactement le même dans le cadre continu. Pour pouvoir exprimer l'inégalité de Talagrand en termes d'influences géométriques, il faut une borne seulement en terme des normes L^1 des dérivées partielles $\partial_i f$, ce qui est fait dans [CE-L]. Les auteurs prouvent ainsi l'inégalité suivante, pour $d\mu(x) = \frac{1}{Z_\beta}e^{-|x|^\beta}dx$, avec $\beta \geq 2$:

$$\mathrm{Var}_\mu(f) \leq C(\rho)\sum_{i=1}^n \frac{\|\partial_i f\|_1(1+\|\partial_i f\|_1)}{(1+\log^+(1/\|\partial_i f\|_1))^{1/2}}$$

et retrouvent les résultats de [K-M-S1] (mais seulement pour $\beta \geq 2$).

Le critère de Benjami-Kalai-Schramm

Nous en venons à la présentation de l'extension du critère de Benjamini, Kalai et Schramm sur des modèles plus généraux et de sa version quantitative. Peu après notre travail, un article de Forströmm est paru recoupant certains de nos résultats, avec toutefois des versions quantitatives plus faibles que les nôtres. Une condition commune qui ressort toutefois est que le critère fonctionne lorsque les constantes d'hypercontractivité ρ et de trou spectral λ sont du même ordre. Dans [For], l'auteur applique ses résultats à la notion "d'exclusion sensitivity". En appendice, nous utilisons nos résultats pour redémontrer un théorème dû à Broman, Garban et Steif [B-G-S].

Les outils de la preuve de [For] sont très proches de ceux d'O'Donnell et Wimmer. L'auteur reprend l'équivalence remarquée par Benjamini, Kalai et Schramm :

$$\mathcal{S}_\rho(f_n) \to 0 \iff \forall \Lambda_n > 0,\ \sum_{\lambda_{i,n}\leq\Lambda_n}\|\widehat{f_{n,i}}\|_2^2 \to 0,$$

où l'on écrit $\{0=\lambda_{0,n}, \lambda_{1,n}, \ldots, \lambda_{i,n}, \ldots\}$, le spectre de L_n (avec $\lambda_{1,n} = \lambda_n$ le trou spectral), et où $f_{n,i}$ sont les fonctions propres associées à $\lambda_{1,n}$. L'auteur borne les sommes à droite associant hypercontractivité et inégalité d'Hölder sous la condition que $(f_n)_{n\geq 0}$ est une suite de fonctions booléennes.

Une condition qui apparaît alors nécéssaire est que $\inf_{n\in\mathbb{N}} \frac{\rho_n}{\lambda_n} > 0$. Comme rappelé dans le cas du cube biaisé $\{0,1\}^n, (p_n\delta_0 + (1-p_n)\delta_1)^{\otimes n}$, le critère B-K-S n'a pas lieu même qualitativement.

Notre preuve s'appuie sur la décomposition de la variance le long du semi-groupe, déjà utilisée dans [CE-L] :

$$\mathrm{Var}_\mu(P_t f) = 2\int_t^\infty \mathscr{E}(P_u f, P_u f)du = 2\sum_{s\in S}\int_t^\infty \|D_s P_u f\|_2^2 du.$$

Dans le cas de graphes de Schreier, la commutation exacte ne permet pas d'intégrer jusqu'à l'infini. Cependant, l'inégalité de trou spectral implique l'inégalité suivante :

$$\forall T > 0,\ \mathrm{Var}_\mu(f) \leq \frac{1}{1-e^{-\lambda T}}\int_0^T \mathscr{E}(P_u f, P_u f)du.$$

Comme dans [CE-L], on utilise ensuite la condition de commutation (4.12). A la différence de l'inégalité de Talagrand, le résultat voulu ne fait intervenir que des normes L^1. Rappelons que nous devons majorer cette quantité par la somme des carrés des influences, c'est-à-dire de la somme des carrés des normes L^1 de $D_s f$. L'hypercontractivité ne permettant pas d'atteindre la norme L^1 (seulement des normes $L^{q(t)}, 1 < q(t) \leq 2$), l'idée est couper l'intégrale en espace en deux bouts, l'un où $D_s P_u f \leq M\|D_s P_u f\|_1$ (ce qui fait apparaître les carrés des normes L^1) et son complémentaire. Nous utilisons l'hypercontractivité sur le sous-ensemble où $D_s P_u f > M\|D_s P_u f\|_1$ pour retrouver la variance de f en utilisant de nouveau sa représentation le long du semi-groupe. Dans le cas de graphes de Schreier (ou Cayley), le résultat obtenu est le suivant :

$$\forall f\ G \to \mathbb{R},\ \mathrm{Var}_\mu(P_t f) \leq \frac{7}{\lambda}\left(\frac{1}{|S|}\sum_{s\in S} I_s(f)^2\right)^{c(1-e^{\rho t})} (\mathrm{Var}_\mu f)^{1-c(1-e^{\rho t})}.$$

Cette estimation quantitative est plus précise que celle établie par Förstromm, et comme dans le travail de Cordero-Erausquin et Ledoux, il n'y a plus besoin d'imposer de conditions particulières sur f (le fait qu'elle soit dans L^2 est déjà nécessaire pour définir sa stabilité au bruit). Notons que nous retrouvons la condition $\inf_{n\in\mathbb{N}} \frac{\rho_n}{\lambda_n} > 0$ pour ce que ce critère soit utilisable. C'est le cas des structures produits (où le rapport est indépendant de n) et des tranches du cube d'ordre k s'il existe $c > 0$ tel que $cn \leq k(1-c)n$. En revanche, ce n'est pas le cas sur le groupe symétrique $\mathfrak{S}_n$. On peut à l'instar des travaux de O'Donnell et Wimmer sur le théorème KKL, se demander si le critère de Benjamini, Kalai et Schramm est satisfait. Par exemple, la suite des fonctions caractéristiques de dérangements f_n satisfait :

$$\frac{1}{n}\sum_{1\leq i > j\leq n} I_{\tau_{ij}}^2(f_n) \leq \frac{4}{n}.$$

On peut se demander si néanmoins, la suite $(f_n)_{n\geq 0}$ est une suite stable au bruit.

Un autre avantage est la souplesse de la preuve permet de l'appliquer aux mêmes modèles que [CE-L], et donc en particulier également à un cadre continu. Pour des produits de mesures log-concaves (par rapport à la gaussienne) à densités régulières (c'est-à-dire tels que $(\mathbb{R},\mu)$ satisfasse $CD(1,\infty)$), le résultat obtenu, exprimé sur les ensembles en termes d'influences géométriques, prend la forme suivante :

$$\mathscr{S}_\eta(A_n) \leq 4\sqrt{1-\eta^2}\left(\sum_{i=1}^n I_i^{\mathscr{G}}(A_n)^2\right)^{c\eta} \mu(A_n)^{2-2c\eta}.$$

Ce résultat est non trivial. En effet, prenons un exemple issu de [K-M-S1]. Considérons l'espace gaussien $(\mathbb{R}^n,\gamma)$. Alors, d'après [K-M-S1] (Proposition 4.3), pour tout $i \in \{1,\dots,n\}$, si l'on note $A_n = (-\infty, a_n]^n$ où a_n est tel que $\gamma((-\infty,a_n])^n = 1/2$, on a :

$$I_i^{\mathscr{G}}(A_n) \approx \frac{(\log n)^{1/2}}{n}.$$

Ainsi, alors que $\gamma(A_n)^2 = 1/4$, $\sum_{i=1}^n I_i^{\mathscr{G}}(A_n)^2 \approx \frac{(\log n)}{n}$. On en tire que $\mathscr{S}_\eta(A_n) \leq 4\left(\frac{\log n}{n}\right)^{c\eta}$, c'est-à-dire que la suite d'ensembles $(A_n)_{n\geq 0}$ est sensible au bruit.

Une autre remarque, issue de [CE-L], est l'implication d'une majoration non triviale de la variance du semi-groupe de la chaleur sur les sphères euclidiennes $\mathbb{S}^n$.