

Grandes déviations et découplage asymptotique convexe pour les champs de variables aléatoires sur $\mathbb{Z}^d$

Table des matières

Introduction	2
1 Cadre	4
1.1 Découplage asymptotique convexe	4
1.2 Contrôle local	5
2 Lemme sous-additif	6
2.1 Énoncé	6
2.2 Démonstration	6
3 PGD faible	9
3.1 Topologie adéquate	9
3.2 PGD	10
3.3 Applications	12
4 Pression	13
4.1 Définition	13
4.2 Pression et entropie	15
5 Annexes	17
5.1 Résultats généraux de grandes déviations	17
5.2 Condition suffisante de PGD faible	17
5.3 Mesures découplées	19
Bibliographie	20

Introduction

Au cours des trente-cinq dernières années, la théorie des grandes déviations pour les moyennes de variables aléatoires a été fortement généralisée, et ce dans les deux directions originelles : les résultats concernant la convergence des moyennes empiriques, d'une part, voie ouverte par Cramér (1938) et Chernoff (1952) ; et, d'autre part, les résultats pour les mesures empiriques, de type Sanov (1957), ou à des niveaux supérieurs (cf. *infra*).

Landford [Lan73] importe le concept de sous-additivité, développé avec Ruelle (1965) en mécanique statistique : il obtient ainsi un principe de grandes déviations (PGD) en dimension finie pour des sommes de variables i.i.d. bornées. L'article de Bahadur et Zabell [BZ79] généralise la théorie de Landford à des variables i.i.d. non nécessairement bornées et à valeurs dans un espace vectoriel localement convexe X de dimension quelconque vérifiant trois hypothèses (notamment la régularité et la convexe-régularité des mesures) : le cadre agréable où ces hypothèses sont vérifiées est celui d'un espace polonais. Ils englobent ainsi, dans leur théorie, les théorèmes de Cramér et de Sanov pour les variables i.i.d. Trois textes développent ces idées par la suite, et présentent des résultats généraux de grandes déviations : les conférences d'Azencott à l'École d'été de probabilités de Saint-Flour [Aze80] et les deux ouvrages de Deuschel et Strook [DS89], d'abord, et de Dembo et Zeitouni [DZ98] ensuite.

Quid des variables dépendantes ? Le théorème de Gärtner-Ellis, énoncé dans [DZ98] est une première tentative : il assure un PGD dans le cas où les pressions partielles

$$p_n(\lambda) = \frac{1}{n} \log \mathbb{E}[\exp\langle \lambda | S_n \rangle]$$

convergent vers une fonction p telle que $\{p < +\infty\}$ contienne 0. L'autre champ d'étude est celui des mesures de Gibbs sur le réseau $\mathbb{Z}^d$. L'ouvrage d'Ellis [Ell85] va dans cette direction et distingue alors trois « niveaux » de PGD :

Niveau 1 : la suite des moyennes empiriques

$$M_n = \frac{1}{|\Lambda(n)|} \sum_{z \in \Lambda(n)} \sigma(z)$$

satisfait un PGD sur X (résultat de type **Cramér**).

Niveau 2 : la suite des mesures empiriques

$$L_n = \frac{1}{|\Lambda(n)|} \sum_{z \in \Lambda(n)} \delta_{\sigma(z)}$$

satisfait un PGD sur $\mathcal{M}_1^+(X)$ (résultat de type **Sanov**).

Niveau 3 : la suite des mesures¹

$$T_n = \frac{1}{|\Lambda(n)|} \sum_{z \in \Lambda(n)} \delta_{\theta_z \omega}$$

satisfait un PGD sur $\mathcal{M}_1^+(X^{\mathbb{Z}^d})$.

Il s'avère que le niveau 3 est le bon cadre pour obtenir le PGD pour les mesures de Gibbs. Les niveaux 1 et 2 sont alors obtenus via le principe de contraction, dans les bons cas. Trois articles sont déposés quasiment simultanément (1986). Comets [Com86] ??? Sans avoir connaissance des travaux de Comets, Olla [Oll88] démontre un PGD pour des sommes de variables i.i.d. dans le cas où X est polonais : il en déduit, via le théorème de Varadhan, un PGD pour les mesures de Gibbs dans le cas où X est polonais et compact. Il retrouve ainsi, par contraction, les résultats de Lanford [Lan73] pour les observables à portée finie. Enfin, Föllmer et Orey [FO88], qui ont vu les prépublications des deux articles précédents, s'attachent au cas où X est fini et proposent une preuve directe du PGD pour les mesures de Gibbs.

Plus tard, Georgii [Geo93] introduit un cadre un peu plus général : il travaille sur un espace de Borel standard² X et introduit la topologie plus naturelle de la convergence locale sur $\mathcal{M}_1^+(X^{\mathbb{Z}^d})$ (alors que les articles précédents considèrent la topologie faible). Il énonce le PGD pour les mesures produits et en déduit, par le théorème de Varadhan, celui pour les mesures de Gibbs (et même une classe plus large de mesures). Pfister [Pfi00] reprend ce cadre topologique et introduit la notion de mesure asymptotiquement découplée, qui permet une généralisation naturelle du lemme sous-additif de Lanford. Il obtient ainsi un PGD de niveau 3 pour ces mesures asymptotiquement découplées.

On voit ainsi réapparaître le fameux lemme sous-additif, pierre angulaire des résultats de type Cramér. Pourtant, les PGD de niveau 3 ne semblent pas pouvoir redonner (même par contraction) le théorème de Cramér pour des variables non bornées. Inspiré des idées de Pfister [Pfi00] et du cours de Cerf [Cer07], qui, travaillant sur les hypothèses topologiques pertinentes de X , déduit le théorème de Sanov du théorème de Cramér, le présent texte propose un cadre unifiant les trois niveaux de PGD : tous sont interprétés au niveau 1.

Les sections 1 à 3 définissent le cadre agréable pour obtenir le PGD annoncé (proposition 3.2.1). On s'intéresse ensuite, dans la partie 4, à relier l'entropie s à la pression p , en particulier à voir à quelles conditions

$$s = -p^*$$

¹Olla [Oll88] et Georgii [Geo93] travaillent plutôt sur

$$\tilde{T}_\Lambda(\omega) = \frac{1}{|\Lambda|} \sum_{z \in \Lambda} \delta_{\theta_z \omega^\Lambda}$$

où ω^Λ est obtenu par périodisation de la boîte Λ . Les résultats obtenus sont analogues.

²Cette hypothèse permet de montrer la coercivité de l'entropie (en utilisant le théorème de Kolmogorov).

1 Cadre

Soit Y est un espace vectoriel localement convexe (evlc) muni d'une tribu $\mathcal{B}$; on note $\mathcal{C}(Y)$ l'ensemble des ouverts convexes mesurables de Y . On suppose que

- (H1) les opérations d'espace vectoriel sont mesurables (pour les tribus produits);
- (H2) $\mathcal{C}(Y)$ est une base d'ouverts.

Si $V \in \mathcal{C}(Y)$ est un voisinage de 0, on note M_V la jauge (ou fonctionnelle de Minkowski) du convexe V .

On considère ici des configurations sur le réseau $\mathbb{Z}^d$

$$\eta : \Omega \rightarrow Y^{\mathbb{Z}^d}$$

invariantes par translation. Rappelons aussi la notation

$$\Lambda(n) = [-n, n]^d \cap \mathbb{Z}^d$$

Cela étant dit, donnons-nous, dans cette première partie, une configuration $\eta : \Omega \rightarrow Y^{\mathbb{Z}^d}$ invariante par translation.

1.1 Découplage asymptotique convexe

La définition fondamentale suivante est reprise de [Pfi00].

Définition 1.1.1 (Découplage asymptotique convexe) *On se donne deux applications g et $c : \mathbb{N} \rightarrow [0, +\infty[$ telles que*

$$\frac{g(n)}{n} \rightarrow 0 \quad \text{et} \quad \frac{c(n)}{|\Lambda(n)|} \rightarrow 0$$

On dit que η est asymptotiquement découplée inférieurement pour les convexes (c.a.d.i.) de paramètres g et c si, pour toutes boîtes finies Λ_1 et Λ_2 , pour tous $C_1 \in \mathcal{C}(Y^{\Lambda_1})$ et $C_2 \in \mathcal{C}(Y^{\Lambda_2})$,

$$\left. \begin{array}{l} \Lambda_1 \subset \Lambda(n) \\ d(\Lambda_2, \Lambda(n)) > g(n) \end{array} \right\} \Rightarrow \mathbb{P}(\eta_{\Lambda_1 \cup \Lambda_2} \in C_1 \times C_2) \geq e^{-c(n)} \mathbb{P}(\eta_{\Lambda_1} \in C_1) \mathbb{P}(\eta_{\Lambda_2} \in C_2)$$

Remarque 1 : On a ainsi une condition de dépendance faible entre ce qui se passe dans une boîte carrée et loin de cette boîte. Cette hypothèse est **globale**. De plus, ce n'est pas seulement la distance entre les deux ensembles de sites supports des événements qui joue, mais la forme de ces ensembles (un cube et l'extérieur de ce cube).

Remarque 2 : Le cas $g = 0$ contient notamment celui des mesures vérifiant l'inégalité FKG.

Remarque 3 : Dans de nombreux cas, l'inégalité est en fait vérifiée pour tous les mesurables, et pas seulement les ouverts convexes.

La proposition suivante est utile et sa démonstration est immédiate.

Proposition 1.1.2 (Stabilité) *Soit $\eta : \Omega \rightarrow Y^{\mathbb{Z}^d}$ une configuration c.a.d.i.. Alors*

- si $y \in Y$, $\eta - y$ est aussi c.a.d.i.;
- si $f : Y \rightarrow Z$ est linéaire, alors $(f(\eta(z)))_{z \in \mathbb{Z}^d}$ est c.a.d.i. également.

1.2 Contrôle local

Définissons aussi une hypothèse pour contrôler chaque site conditionnellement au reste de la configuration.

Définition 1.2.1 (Contrôle local) *On dit que η est contrôlée localement si, pour tout $V \in \mathcal{C}(Y)$ voisinage de 0, il existe $t > 0$ et $\alpha > 0$ tels que, pour toute boîte $\Lambda \subset \mathbb{Z}^d \setminus \{0\}$ et tout convexe $C \in \mathcal{C}(Y^\Lambda)$,*

$$\mathbb{P}(\eta(0) \in tV | \eta_\Lambda \in C) \geq \alpha$$

Remarque : Il est important ici encore de travailler avec des convexes. En effet, on utilisera, par la suite, le fait que pour tout $t' \geq t$,

$$\mathbb{P}(\eta(0) \in t'V | \eta_\Lambda \in C) \geq \alpha$$

ce qui est vérifié car $tV \subset t'V$.

Proposition 1.2.2 (Stabilité) *Soit $\eta : \Omega \rightarrow Y^{\mathbb{Z}^d}$ une configuration contrôlée localement. Alors*

- si $y \in Y$, $\eta - y$ est aussi contrôlée localement ;
- si $f : Y \rightarrow Z$ est linéaire, alors $(f(\eta(z)))_{z \in \mathbb{Z}^d}$ est aussi contrôlée localement.

Exemples : Voici quelques exemples de configurations contrôlées localement :

- η faiblement borné.
En particulier, $\eta(z) = \theta_z \delta_\omega$ (**niveau 3**, cadre traité dans [Pfi00]) ; en effet, les voisinages de 0 pour la τ -topologie absorbent tous la boule unité de $\mathcal{F}_{loc}^*$, donc *a fortiori* $\mathcal{M}_1^+$.
- η c.a.d.i. avec $g(0) = 0$.
En particulier, η vérifiant FKG, η i.i.d. (**niveau 1**).

2 Lemme sous-additif

2.1 Énoncé

La notion de **v.a. c.a.d.i. invariante par translation** est une généralisation naturelle de celle de **suite de v.a. i.i.d.**. En lui adjoignant l'hypothèse de contrôle local, on conserve une forme de sous-additivité. En effet, introduisons la notation

$$m_\Lambda \eta = \frac{1}{|\Lambda|} \sum_{z \in \Lambda} \eta(z)$$

Donnons-nous deux ouverts convexes de la forme

$$C = y + V \quad \text{et} \quad C(\varepsilon) = y + (1 - \varepsilon)V$$

où $V \in \mathcal{C}(Y)$ est un voisinage de 0 et $\varepsilon \in]0, 1[$. Prenons aussi deux marges de manœuvre $0 < \delta_1 < 1$ et $\delta_2 > 0$.

Théorème 2.1.1 (lemme sous-additif) *Supposons η c.a.d.i., de loi invariante par translation et contrôlée localement.*

Alors,

$$\begin{aligned} \exists M(\varepsilon, \delta_1, \delta_2) \quad \forall m \geq M \quad \exists N(m, \varepsilon, \delta_1, \delta_2) \quad \forall n \geq N \\ \frac{1}{|\Lambda(n)|} \log \mathbb{P}(m_{\Lambda(n)} \eta \in C) \geq (1 - \delta_1) \frac{1}{|\Lambda(m)|} \log \mathbb{P}(m_{\Lambda(m)} \eta \in C(\varepsilon)) - \frac{c(m)}{|\Lambda(m)|} - \delta_2 \end{aligned}$$

2.2 Démonstration

La démonstration suit les étapes de celle de Pfister [Pfi00] en l'adaptant à ce nouveau cadre.

Si m et $n \in \mathbb{N}^*$, la division euclidienne de n par $m + g(m)$ s'écrit

$$n = k(m + g(m)) + r$$

On pave alors $\Lambda(n)$ avec k^d boîtes isométriques à $\Lambda(m + g(m))$, que l'on note Λ'_q , pour $q \in [1, k^d]$. Puis, on définit Λ_q comme étant la boîte isométrique à $\Lambda(m)$ centrée dans Λ'_q ; et, enfin,

$$\Lambda_{k^d+1} = \Lambda(n) \setminus \bigcup_{q=1}^{k^d} \Lambda_q$$

est l'ensemble des sites marginaux. Pour justifier cette dénomination, notons que,

$$\rho_{m,n} = \frac{|\Lambda_{k^d+1}|}{|\Lambda(n)|} \xrightarrow[n \geq m]{m \rightarrow \infty} 0$$

Rappelons que la jauge du convexe V permet d'écrire

$$m_{\Lambda(n)}\eta \in C \iff M_V(m_{\Lambda(n)}\eta - y) < 1$$

Alors, en notant ϕ la fonction convexe $M_V(\cdot - y)$,

$$\phi(m_{\Lambda(n)}\eta) \leq \sum_{q=1}^{k^d} \frac{1}{k^d} \phi(m_{\Lambda_q}\eta) + \frac{|\Lambda_{k^d+1}|}{|\Lambda(n)|} \phi(m_{\Lambda_{k^d+1}}\eta)$$

D'où

$$\frac{1}{|\Lambda(n)|} \log \mathbb{P}(m_{\Lambda(n)}\eta \in C) = \frac{1}{|\Lambda(n)|} \log \mathbb{P}(\phi(m_{\Lambda(n)}\eta) < 1) \geq P_{m,n} + R_{m,n} \quad (1)$$

où

$$P_{m,n} = \frac{1}{|\Lambda(n)|} \log \mathbb{P}\left(\bigcap_{q=1}^{k^d} \{m_{\Lambda_q}\eta \in C(\varepsilon)\}\right)$$

est le terme principal, et

$$R_{m,n} = \frac{1}{|\Lambda(n)|} \log \mathbb{P}\left(\rho_{m,n} \phi(m_{\Lambda_{k^d+1}}\eta) < \varepsilon \mid \sum_{q=1}^{k^d} \frac{1}{k^d} \phi(m_{\Lambda_q}\eta) < 1 - \varepsilon\right)$$

le reste.

Terme principal : On traite le premier terme grâce au découplage asymptotique de η qui permet d'écrire :

$$\begin{aligned} \frac{1}{|\Lambda(n)|} \log \mathbb{P}\left(\bigcap_{q=1}^{k^d} \{m_{\Lambda_q}\eta \in C(\varepsilon)\}\right) &\geq \frac{1 - \rho_{m,n}}{k^d |\Lambda(m)|} \log \left((e^{-c(m)})^{k^d-1} \prod_{q=1}^{k^d} \mathbb{P}(m_{\Lambda_q}\eta \in C(\varepsilon)) \right) \\ &\geq (1 - \delta_1) \frac{1}{|\Lambda(m)|} \log \mathbb{P}(m_{\Lambda(m)}\eta \in C(\varepsilon)) - \frac{c(m)}{|\Lambda(m)|} \end{aligned}$$

quitte à choisir $m \geq M(\delta_1)$ et $n \geq N(m, \delta_1)$.

Terme résiduel : Il s'agit de s'assurer que, quitte à choisir $m \leq n$ suffisamment grands, on peut rendre $R_{m,n}$ arbitrairement proche de 0. Tirons profit du contrôle local. L'hypothèse pour $\eta - y$ donne un $t > 0$ et un $\alpha > 0$. On choisit $m \geq M(\varepsilon)$ et $n \geq N(m, \varepsilon)$ de sorte que

$$\frac{\varepsilon}{\rho_{m,n}} > t$$

Dès lors, on obtient par récurrence :

$$\begin{aligned}
R_{m,n} &\geq \frac{1}{|\Lambda(n)|} \log \frac{\mathbb{P} \left(\forall z \in \Lambda_{k^d+1} \quad \rho_{m,n} \phi(\eta_z) < \varepsilon \quad ; \quad \sum_{q=1}^{k^d} \frac{1}{k^d} \phi(m_{\Lambda_q} \eta) < 1 - \varepsilon \right)}{\mathbb{P} \left(\sum_{q=1}^{k^d} \frac{1}{k^d} \phi(m_{\Lambda_q} \eta) < 1 - \varepsilon \right)} \\
&\geq \frac{|\Lambda_{k^d+1}|}{|\Lambda(n)|} \log \alpha
\end{aligned}$$

et ainsi $R_{m,n} \geq -\delta_2$ quitte à prendre, au final, $m \geq M(\varepsilon, \delta_2)$ et $n \geq N(m, \varepsilon, \delta_2)$.

□

Remarque 1 : On a l'ordre de grandeur suivant :

$$\rho_{m,n} \lesssim d \cdot \left(\frac{g(m)}{2m + g(m)} + \frac{r}{2n} \right)$$

Ainsi, pour que $\rho_{m,n}$ puisse être rendu arbitrairement petit, il faut que le premier terme tende vers 0 avec m , ce qui équivaut à

$$\frac{g(m)}{m} \rightarrow 0$$

L'hypothèse mise sur g est donc ici justifiée.

Remarque 2 : Intuitivement, on a d'abord fait en sorte que l'écartement entre les petites boîtes soit suffisamment petit devant la taille de ces boîtes. Puis, on a pris suffisamment de petites boîtes pour que la proportion occupée par elles dans une grande boîte soit assez grande. Au final, on obtient une sous-additivité asymptotique entre les petites boîtes et la grande.

3 PGD faible

On pourrait dès maintenant énoncer un principe de grandes déviations pour une variable η c.a.d.i. contrôlée localement, mais on se place dans un cadre plus général, qui ne complexifie pas la démonstration et qui s'applique dans de nombreux autres cas (notamment, $\sigma(z) = \delta_{\theta_z \omega}$, qui n'est pas c.a.d.i.).

On se donne donc $\sigma : \Omega \rightarrow X^{\mathbb{Z}^d}$ de loi invariante par translation, mais pas nécessairement c.a.d.i.

3.1 Topologie adéquate

Comme annoncé, le découplage asymptotique n'est pas nécessaire pour qu'il y ait PGD : le tout est d'exploiter une topologie idoine sur X , que l'on va construire.

À cet effet, on suppose donnée une famille

$$\mathcal{Y} = (Y_i, f_i)_{i \in I}$$

telle que

(P1) pour tout $i \in I$, (Y_i, τ_i) est un evlc, $f_i : X \rightarrow Y_i$ est linéaire ;

(P2) pour tous $i, j \in I$,

$$(Y_i \times Y_j, (f_i, f_j)) \in \mathcal{Y}$$

Par convention, si $i = (i_1, \dots, i_p)$, on notera

$$(Y_i, f_i) = (Y_{i_1} \times \dots \times Y_{i_p}, (f_{i_1}, \dots, f_{i_p}))$$

(P3) pour tout $i \in I$,

$$\eta_i = (f_i(\sigma_z))_{z \in \mathbb{Z}^d}$$

est c.a.d.i. et contrôlée localement.

On munit alors X de la topologie initiale pour la famille $(Y_i, f_i)_{i \in I}$, topologie d'evlc (cf. [Bou81]) que l'on notera τ et dont on désignera par $\mathcal{C}$ la base des ouverts convexes de la forme

$$A = \bigcap_{j \in J} \{x \in X; f_j(x) \in C_j\} = \{x \in X; f_i(x) \in C_i\}$$

avec $J = \{i_1, \dots, i_r\}$ sous-ensemble fini de I , $C_j \in \mathcal{C}(Y_j)$ et $i = (i_1, \dots, i_r)$.

Description équivalente : On se donne un système projectif

$$\mathcal{Y} = (Y_i, (\pi_{ji})_{j \leq i})_{i \in I}$$

tel que

(P1') pour tout $i \in I$, (Y_i, τ_i) est un evlc et la projection $\pi_{ji} : Y_i \rightarrow Y_j$ est linéaire ;

(P2') $X = \text{proj lim } Y_i$ et $f_i : X \rightarrow Y_i$ est la projection canonique sur Y_i ;

(P3') pour tout $i \in I$,

$$\eta_i = (f_i(\sigma_z))_{z \in \mathbb{Z}^d}$$

est c.a.d.i. et contrôlée localement.

Exemples :

- si X est déjà un evlc et si $\mathcal{Y}$ est engendré par $f_0 = Id$ ($Y_0 = X$), on récupère la topologie sur X ;
- si X est un evlc et si $\mathcal{Y}$ est engendré par X' ($Y_i = \mathbb{R}$ pour $i \in X'$), on tombe sur la topologie $\sigma(X, X')$;
- pour X' dual d'un evlc, si $\mathcal{Y}$ est engendrée par X ($Y_i = \mathbb{R}$ pour $i \in X$), on a la topologie $\sigma(X', X)$.

3.2 PGD

Le décor est maintenant planté. On peut énoncer :

Proposition 3.2.1 (PGD) Si $\sigma : \Omega \rightarrow X^{\mathbb{Z}^d}$ a les propriétés (P1), (P2) et (P3), alors

$$\left(\frac{1}{|\Lambda(n)|} \sum_{z \in \Lambda(n)} \sigma(z) \right)_{n \geq 0}$$

vérifie un PGD faible gouverné par

$$s(x) = \inf_{\substack{A \in \mathcal{C} \\ A \ni x}} \liminf_{n \rightarrow \infty} \frac{1}{|\Lambda(n)|} \log \mathbb{P}(m_{\Lambda(n)} \sigma \in A)$$

Remarque : En vertu de (2),

$$s(x) = \inf_{\substack{A \in \mathcal{C} \\ A \ni x}} s(A) = \inf_{\substack{A \in \mathcal{C} \\ A \ni x}} \bar{s}(A)$$

Démonstration : Donnons-nous $A \in \mathcal{C}$; il est de la forme

$$A = x + U$$

avec

$$U = \{x \in X; f(x) \in V\}$$

voisinage de 0. Définissons

$$C = f(x) + V$$

puis

$$A_\varepsilon = x + (1 - \varepsilon)U \quad \text{et} \quad C(\varepsilon) = f(x) + (1 - \varepsilon)V$$

La famille $\eta = (f(\sigma_z))_{z \in \mathbb{Z}^d}$ vérifie les hypothèses du lemme sous-additif 2.1.1 : on en déduit, en faisant tendre n , puis m , vers l'infini, et enfin δ_1 et δ_2 vers 0^+ ,

$$\liminf \frac{1}{|\Lambda(n)|} \log \mathbb{P}(m_{\Lambda(n)} \eta \in C) \geq \limsup \frac{1}{|\Lambda(m)|} \log \mathbb{P}(m_{\Lambda(m)} \eta \in C(\varepsilon))$$

Pour toute boîte finie Λ , la linéarité de f assure

$$\{m_\Lambda \sigma \in A\} = \{f(m_\Lambda \sigma) \in C\} = \{m_\Lambda \eta \in C\}$$

et l'égalité analogue avec des ε en indice. On en déduit

$$\underline{s}(A) \geq \bar{s}(A_\varepsilon)$$

On conclut alors grâce au lemme 5.2.1. □

En termes de **limite projective**, on peut penser exploiter le théorème de Dawson-Gärtner (cf. [DZ98]) : on obtient alors un PGD pour σ , à condition que les s_η soient de bonnes fonctions de taux, ce qui est moins bon que le résultat obtenu. Il n'est pas étonnant qu'on n'obtienne pas ainsi un résultat optimal étant donné que le théorème de Dawson-Gärtner ne se sert pas de la linéarité des f_i .

Proposition 3.2.2 *L'entropie s est s.c.s. et concave.*

Démonstration : Elle est adaptée de [Cer07].

La semi-continuité est immédiate : si $t \in \mathbb{R}$ et $x \in \{s < t\}$, il existe $A \in \mathcal{C}$ contenant x tel que $\underline{s}(A) < t$. Alors, pour tout $y \in A$, $s(y) \leq \underline{s}(A) < t$, donc $x \in A \subset \{s < t\}$ et $\{s < t\}$ est ouvert.

Pour la concavité, il suffit alors de prouver que

$$s\left(\frac{x+y}{2}\right) \geq \frac{s(x) + s(y)}{2}$$

La semi-continuité permet alors de passer des inégalités

$$s((1-t)x + ty) \geq (1-t)s(x) + ts(y)$$

pour t dyadique à celles pour $t \in [0, 1]$.

La preuve est un raffinement de la démonstration du lemme sous-additif (2.1.1). Soit $A \in \mathcal{C}$ contenant $\frac{x+y}{2}$. Par continuité des applications d'espace vectoriel, il existe $\varepsilon \in]0, 1[$ tel que $A(\varepsilon) \ni \frac{x+y}{2}$, puis A_x et $A_y \in \mathcal{C}$, contenant respectivement x et y , tels que $\frac{1}{2}A_x + \frac{1}{2}A_y \subset A(\varepsilon)$. Soit enfin $\delta_1 \in]0, 1[$ et $\delta_2 > 0$. On se restreint, dans le calcul suivant, aux couples (m, n) tels que k est pair.

$$\begin{aligned} & \frac{1}{|\Lambda(n)|} \log \mathbb{P}(m_{\Lambda(n)} \sigma \in A) \\ & \geq \frac{1}{|\Lambda(n)|} \log \mathbb{P}(m_{\Lambda(n)} \sigma \in A(\varepsilon)) + R_{m,n} \\ & \geq \frac{1}{|\Lambda(n)|} \log \mathbb{P}\left(\bigcap_{q \text{ pair}} \{m_{\Lambda_q} \sigma \in A_x\} \cap \bigcap_{q \text{ impair}} \{m_{\Lambda_q} \sigma \in A_y\}\right) + R_{m,n} \end{aligned}$$

En poursuivant comme précédemment, on obtient

$$\underline{s}(A) \geq \frac{1}{2}(\underline{s}(A_x) + \underline{s}(A_y)) \geq \frac{1}{2}(s(x) + s(y))$$

puis l'inégalité voulue en passant à l'infimum sur les $A \ni \frac{x+y}{2}$.

3.3 Applications

Voici un certain nombre de cas où il y a PGD :

- pour les suites de v.a. i.i.d. ($g = 0$ et $d = 1$) (théorème de Cramér i.i.d.) ;
- pour les configurations vérifiant l'inégalité FKG ($g = 0$), notamment le modèle d'Ising et, plus généralement, les mesures de Gibbs à potentiel borné ;
- pour les variables σ c.a.d.i. de loi invariante par translation et faiblement bornées, notamment le modèle d'Ising et les mesures de Gibbs à potentiel borné ;
- pour les mesures empirique de σ c.a.d.i. de loi invariante par translation (théorème de Sanov).

La question est maintenant de savoir si l'on peut, comme dans le cas i.i.d., définir une pression p et si

$$s = -p^*$$

moyennant quelques hypothèses. Même dans le cas i.i.d., ce n'est pas toujours le cas (cf. [Cer07]).

4 Pression

4.1 Définition

On définit, pour toute boîte $\Lambda \subset \mathbb{Z}^d$,

$$\forall \lambda \in X^* \quad p_\Lambda(\lambda) = \frac{1}{|\Lambda|} \log \mathbb{E} \left[\exp \left\langle \lambda \left| \sum_{z \in \Lambda} \sigma(z) \right. \right\rangle \right]$$

Notons que X^* désigne le dual topologique de X ; on suppose³ dans la suite que

$$X^* = \{\lambda_i \circ f_i; i \in I, \lambda_i \in Y_i^*\}$$

En particulier, pour tout $\lambda \in X^*$, la configuration

$$\eta = (\langle \lambda | \sigma(z) \rangle)_{z \in \mathbb{Z}^d} : \Omega \rightarrow \mathbb{R}^{\mathbb{Z}^d}$$

est c.a.d.i. et contrôlée localement (cf. 1.1.2 et 1.2.2).

On veut maintenant définir la pression p en volume infini. Tout se passe à merveille ainsi que le montre la proposition suivante :

Proposition 4.1.1 (Existence de p) *Pour tout $\lambda \in X^*$, la suite $(p_{\Lambda(n)}(\lambda))$ converge dans $[-\infty, +\infty]$. On note $p(\lambda)$ sa limite.*

Démonstration : La preuve est analogue à celle du lemme sous-additif 2.1.1.

On définit $\eta(z) = \langle \lambda | \sigma(z) \rangle$ et on écrit

$$p_{\Lambda(n)}(\lambda) = \frac{1}{|\Lambda(n)|} \log \mathbb{E}[P_{m,n} \cdot R_{m,n}]$$

avec, de même, un terme principal

$$P_{m,n} = \prod_{q=1}^{k^d} \exp \left(\sum_{z \in \Lambda_q} \eta(z) \right)$$

³En fait, cela doit être vrai dans de nombreux cas ; l'inclusion $\supset$ est évidente ; voici une ébauche de démonstration de l'autre, inspirée de [Rud73]. Soit $\lambda \in X^*$. Il existe alors un voisinage convexe de 0 V sur lequel $|\lambda| < 1$. Or V est de la forme

$$V = \{x \in X; \forall \mu \in A \quad |\mu \circ f_i(x)| < 1\}$$

avec $A \subset Y_i^*$ (Hahn-Banach). Puis, sachant que $M_V(x) = \sup_{\mu \in A} |\mu \circ f_i(x)|$, on obtient

$$\forall x \in X \quad |\lambda(x)| \leq \sup_{\mu \in A} |\mu \circ f_i(x)|$$

d'où

$$\lambda = g((\mu \circ f_i)_{\mu \in A})$$

avec g linéaire. Faut-il ensuite rajouter l'hypothèse que Y_i^* est complet ?

et le reste

$$R_{m,n} = \mathbb{E} \left[\exp \left(\sum_{z \in \Lambda_{k^d+1}} \eta(z) \right) \middle| \prod_{q=1}^{k^d} \exp \left(\sum_{z \in \Lambda_q} \eta(z) \right) \right]$$

Pour qu'il y ait convergence, il suffit de s'assurer que

– l'on peut minorer le reste

$$R_{m,n} \geq e^{-\delta_2 |\Lambda(n)|}$$

pour m et n assez grands.

– l'hypothèse de découplage s'applique bien pour traiter le terme principal.

Terme résiduel : Encore une fois, on va utiliser le contrôle local.

Lemme 4.1.2 *Il existe $\beta > 0$ tel que, pour toute boîte finie $\Lambda \subset \mathbb{Z}^d \setminus \{0\}$,*

$$\mathbb{E}[e^{\eta(0)} | \eta_\Lambda] \geq \beta$$

Démonstration du lemme : On note que, pour tous Λ finie et $t > 0$,

$$\begin{aligned} \mathbb{E}[e^{\eta(0)} | \eta_\Lambda] &\geq \mathbb{E}[e^{\eta(0)} 1_{|\eta(0)| < t} | \eta_\Lambda] \\ &\geq e^{-t} \mathbb{P}(|\eta(0)| < t | \eta_\Lambda) \end{aligned}$$

L'hypothèse de contrôle local appliqué à η et $V =]-1, 1[$ donne un t et un $\alpha > 0$, de sorte que

$$\begin{aligned} \mathbb{P}(|\eta(0)| < t | \eta_\Lambda)(\omega) &= \lim_{r \rightarrow 0^+} \mathbb{P}(|\eta(0)| < t | \eta_\Lambda \in B(\eta_\Lambda(\omega), r)) \\ &\geq \alpha \end{aligned}$$

L'égalité de la première ligne est due au fait que l'ensemble de Lebesgue d'une fonction de $L^1(\mathbb{R}^\Lambda; \mathbb{R})$ est de mesure pleine (cf. par exemple [Rud66]).

Ainsi, $\beta = e^{-t}\alpha > 0$ convient.

□

On déduit du lemme, par récurrence, que

$$R_{m,n} \geq \beta^{|\Lambda_{k^d+1}|} \geq e^{-\delta_2 |\Lambda(n)|}$$

pour $m \geq M(\delta_2)$ et $n \geq N(m, \delta_2)$.

Terme principal : sachant que les $\eta(z)$ sont des formes linéaires en σ ,

$$\left\{ \exp \sum_{z \in \Lambda_q} \eta(z) > t \right\} \in \mathcal{C}^{|\Lambda_q|}$$

(il s'agit en fait d'un demi-espace ouvert).

Ainsi,

$$\begin{aligned}
& \frac{1}{|\Lambda(n)|} \log \mathbb{E} \left[\prod_{q=1}^{k^d} \exp \left(\sum_{z \in \Lambda_q} \eta(z) \right) \right] \\
&= \frac{1}{|\Lambda(n)|} \log \mathbb{E} \left[\prod_{q=1}^{k^d} \int_{t_q \geq 0} 1_{\exp \left(\sum_{z \in \Lambda_q} \eta(z) \right) > t_q} \right] \\
&= \frac{1}{|\Lambda(n)|} \log \int_{t_1, \dots, t_{k^d} \geq 0} \mathbb{E} \left[\prod_{q=1}^{k^d} 1_{\exp \left(\sum_{z \in \Lambda_q} \eta(z) \right) > t_q} \right] \\
&\geq \frac{1}{|\Lambda(n)|} \log \int_{t_1, \dots, t_{k^d} \geq 0} (e^{-c(m)})^{k^d-1} \prod_{q=1}^{k^d} \mathbb{E} \left[1_{\exp \left(\sum_{z \in \Lambda_q} \eta(z) \right) > t_q} \right] \\
&\geq \frac{k^d}{|\Lambda(n)|} \log \mathbb{E} \left[\exp \left(\sum_{z \in \Lambda_1} \eta(z) \right) \right] - \frac{c(m)}{|\Lambda(m)|}
\end{aligned}$$

On obtient finalement que, pour $m \geq M(\delta_1, \delta_2)$ et $n \geq N(m, \delta_1, \delta_2)$,

$$\begin{aligned}
p_{\Lambda(n)}(\lambda) &\geq \frac{1}{|\Lambda(n)|} \log \mathbb{E} [P_{m,n} \cdot e^{-\delta_2 |\Lambda(n)|}] \\
&\geq (1 - \delta_1) p_{\Lambda(m)}(\lambda) - \frac{c(m)}{|\Lambda(m)|} - \delta_2
\end{aligned}$$

D'où

$$\limsup p_{\Lambda(n)}(\lambda) \geq \liminf p_{\Lambda(m)}(\lambda)$$

autrement dit $p_{\Lambda(n)}(\lambda)$ converge.

□

Proposition 4.1.3 (Convexité de p) *La pression p est convexe.*

Démonstration : l'inégalité de Hölder assure que les p_Λ sont convexes. Puis, sachant qu'une limite supérieure de fonctions convexes est convexe, p l'est également.

4.2 Pression et entropie

L'inégalité de Tchebychev exponentielle (cf. [Cer07]) prend ici la forme suivante :

Proposition 4.2.1 (Inégalité de Tchebycheff exponentielle) *Pour toute boîte $\Lambda \subset \mathbb{Z}^d$ et $A \in \mathcal{B}$,*

$$\mathbb{P}(m_\Lambda \sigma \in A) \leq \exp \left[-|\Lambda| \sup_{\lambda \in X^*} \left(\inf_{a \in A} \langle \lambda | a \rangle - p_\Lambda(\lambda) \right) \right]$$

On en déduit, en considérant $\Lambda = \Lambda(n)$ et en faisant tendre n vers ∞ :

$$\underline{s}(A) \leq \liminf_{\lambda \in X^*} \inf_{\substack{A \in \mathcal{C} \\ A \ni x}} \left(p_{\Lambda(n)}(\lambda) - \inf_{a \in A} \langle \lambda | a \rangle \right) \leq \inf_{\lambda \in X^*} \left(p(\lambda) - \inf_{a \in A} \langle \lambda | a \rangle \right)$$

Puis,

$$\begin{aligned} s(x) &= \inf_{\substack{A \in \mathcal{C} \\ A \ni x}} \underline{s}(A) \\ &\leq \inf_{\substack{A \in \mathcal{C}_w \\ A \ni x}} \underline{s}(A) = \underline{s}_m(x) \\ &\leq \inf_{\substack{A \in \mathcal{C}_w \\ A \ni x}} \inf_{\lambda \in X^*} \left(p(\lambda) - \inf_{a \in A} \langle \lambda | a \rangle \right) \\ &= \inf_{\lambda \in X^*} \inf_{\substack{A \in \mathcal{C}_w \\ A \ni x}} \left(p(\lambda) - \inf_{a \in A} \langle \lambda | a \rangle \right) \\ &= - \sup_{\lambda \in X^*} \left(\langle \lambda | x \rangle - p(\lambda) \right) \\ &= -p^*(x) \end{aligned}$$

Résumons :

Théorème 4.2.2 *On a toujours*

$$s \leq -p^*$$

On voudrait maintenant savoir dans quels cas l'égalité est vérifiée.

Théorème 4.2.3 *Soit μ la loi de $\sigma(0)$. On a l'égalité*

$$s = -p^*$$

dans les deux cas suivants :

1. *$\text{cosupp}(\mu)$ est relativement compact ;*
2. *$\text{cosupp}(\mu)$ est faiblement borné, le PGD est fort et s est coercive.*

Démonstration : Le résultat est conséquence du lemme de Varadhan 5.1.2.

5 Annexes

5.1 Résultats généraux de grandes déviations

Lemme 5.1.1 *Si $u^{(1)}, \dots, u^{(r)}$ sont r suites positives, on a l'égalité*

$$\limsup \frac{1}{n} \log \sum_{k=1}^r u_n^{(k)} = \max_{k \in [1, r]} \limsup \frac{1}{n} \log u_n^{(k)}$$

On trouvera une démonstration de ce résultat dans [Cer07].

Lemme 5.1.2 (Varadhan) *On suppose que (μ_n) vérifie un PGD sur X gouverné par une entropie s . Soit $f : X \rightarrow \mathbb{R}$ s.c.s., mesurable et majorée. Alors, si s est coercive ou si f est minorée (donc bornée),*

$$\frac{1}{n} \log \int_X e^{nf(x)} \mu_n(dx) \rightarrow \sup_X [f + s]$$

Démonstration : La démonstration dans le cas où s est coercive est dans [Cer07]. Dans le cas où f est bornée, seule la démonstration de l'inégalité $\leq$ change. Étant donné $N > 0$, on discrétise

$$X = \bigcup_{k \in J} \underbrace{\left\{ x; f(x) \in \left[\frac{k}{N}, \frac{k+1}{N} \right] \right\}}_{X_k}$$

L'union est indexée par un ensemble fini J , car f est bornée, et les X_k sont fermés, par continuité de f ; on obtient donc, grâce au lemme 5.1.1

$$\begin{aligned} \limsup \frac{1}{n} \log \int_X e^{nf(x)} \mu_n(dx) &\leq \limsup \frac{1}{n} \log \sum_{k \in J} \int_{X_k} e^{nf(x)} \mu_n(dx) \\ &\leq \limsup \frac{1}{n} \log \sum_{k \in J} e^{n \frac{k+1}{N}} \mu(X_k) \\ &\leq \max_{k \in J} \left[\frac{k+1}{N} + \sup_{X_k} s \right] \\ &\leq \max_{k \in J} \left[\sup_{X_k} f + \frac{1}{N} + \sup_{X_k} s \right] \\ &\leq \sup_X [f + s] + \frac{1}{N} \end{aligned}$$

5.2 Condition suffisante de PGD faible

Soit X un evlc. On suppose toujours vérifiées les hypothèses (H1) et (H2) et on note $\mathcal{C}$ la base de ses ouverts convexes mesurables. On suppose donnée une suite (μ_n) de

probabilités sur X : le but de cette partie est de dégager une condition suffisante pour que (μ_n) satisfasse un PGD faible, gouverné nécessairement par

$$s(x) = \inf_{\substack{A \in \mathcal{C} \\ A \ni x}} \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(A) = \inf_{\substack{A \in \mathcal{C} \\ A \ni x}} \liminf_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(A) \quad (2)$$

Définissons, pour $A \in \mathcal{C}$,

$$\underline{s}(A) = \liminf_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(A) \quad \text{et} \quad \bar{s}(A) = \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(A)$$

Lemme 5.2.1 *Si*

$$\forall x \in X \quad \forall A \in \mathcal{C} \text{ avec } x \in A \quad \exists A' \in \mathcal{C} \text{ avec } x \in A' \subset A \quad \underline{s}(A) \geq \bar{s}(A')$$

alors (μ_n) vérifie un PGD faible gouverné par s .

Remarque : Bien que dans le cas des μ_n i.i.d., $\underline{s} = \bar{s}$, l'égalité n'est pas nécessaire pour qu'il y ait PGD (faible). Et, de fait, elle n'est plus vraie dans le cas seulement c.a.d.i. qui nous occupe ici.

Démonstration : C'est un argument général de théorie des grandes déviations. Nous reproduisons ici la présentation de [Cer07].

En passant à l'infimum sur les voisinages dans l'hypothèse du lemme, on obtient l'égalité

$$\boxed{\inf_{\substack{A \in \mathcal{C} \\ A \ni x}} \underline{s}(A) = \inf_{\substack{A \in \mathcal{C} \\ A \ni x}} \bar{s}(A)} \quad (3)$$

C'est juste ce qu'il nous faut pour prouver le PGD faible. En effet, la borne inférieure est conséquence immédiate de l'égalité

$$s(x) = \inf_{\substack{A \in \mathcal{C} \\ A \ni x}} \underline{s}(A)$$

Pour la borne supérieure, soit K un compact mesurable de X et $\delta > 0$. Par définition de s ,

$$\forall x \in K \quad \exists A_x \in \mathcal{C} \quad \bar{s}(A_x) \leq \max(s(x) + \delta, -1/\delta)$$

On extrait alors un sous-recouvrement fini de K par les A_x , noté $\{A_x; x \in J\}$, avec J fini. Puis

$$\begin{aligned} \limsup \frac{1}{n} \log \mu_n(K) &\leq \limsup \frac{1}{n} \log \sum_{x \in J} \mu_n(A_x) \\ &\leq \max_{x \in J} \limsup \frac{1}{n} \log \mu_n(A_x) \\ &= \max_{x \in J} \bar{s}(A_x) \\ &\leq \max_{x \in J} \max(s(x) + \delta, -1/\delta) \\ &\leq \max \left(\max_{x \in K} s(x) + \delta, -1/\delta \right) \end{aligned}$$

ce qui donne la borne inférieure, en faisant tendre δ vers 0^+ .

□

5.3 Mesures découplées

Le modèle d'Ising (et plus généralement les mesures de Gibbs à espace d'états borné) est c.a.d.i. avec $g(0) = 0$. On montre par exemple que, pour $A \in \mathcal{F}_{\mathbb{Z}^d \setminus \{0\}}$,

$$\mathbb{P}(\sigma(0) = 1, A) \geq \gamma \mathbb{P}(\sigma(0) = 1) \mathbb{P}(A)$$

avec

$$\gamma = \frac{1}{(1 + e^{8\beta}) \mathbb{P}(\sigma(0) = 1)} > 0$$

Références

- [Aze80] R. AZENCOTT – *Grandes déviations et applications*, École d'été de probabilités de Saint-Flour, Lecture Notes in Mathematics, vol. 774, Springer-Verlag (1980).
- [Bou81] N. BOURBAKI – *Espaces vectoriels topologiques*, Masson (1981).
- [BZ79] R. R. BAHADUR & S. L. ZABELL – Large Deviations of the Sample Mean in General Vector Spaces, *Ann. Prob.* **7**, no. 4, p. 587–621 (1979).
- [Cer07] R. CERF – *On Cramér's Theory in Infinite Dimensions*, Panoramas et synthèses (2007).
- [Com86] F. COMETS – Grandes déviations pour des champs de Gibbs sur $\mathbb{Z}^d$, *C. R. Acad. Sci. Série 1*, **303**, p. 511–514 (1986).
- [DS89] J.-D. DEUSCHEL & D. W. STROOCK – *Large Deviations*, Academic Press (1989).
- [DZ98] A. DEMBO & O. ZEITOUNI – *Large Deviations Techniques and Applications*, Springer-Verlag (1998).
- [Ell85] R. S. ELLIS – *Large Deviations and Statistical Mechanics*, Springer (1985).
- [FO88] H. FÖLLMER & S. OREY – Large Deviations for the Empirical Field of a Gibbs Measure, *Ann. Prob.* **16**, p. 961–977 (1988).
- [Geo93] H. O. GEORGII – Large Deviations and Maximum Entropy Principle for Interacting Random Fields on $\mathbb{Z}^d$, *Ann. Prob.* **21**, p. 1845–1875 (1993).
- [Lan73] O. E. LANFORD – *Entropy and Equilibrium States in Classical Statistical Mechanics*, Statistical Mechanics and Mathematical Problems, Lecture Notes in Physics, vol. 20, Springer (1973).
- [Oll88] S. OLLA – Large Deviations for Gibbs Random Fields, *Prob. Th. Rel. Fields* **77**, p. 343–357 (1988).
- [Pfi00] C.-E. PFISTER – *Thermodynamical Aspects of Classical Lattice Systems* (2000).
- [Rud66] W. RUDIN – *Real and Complex Analysis*, McGraw-Hill (1966).
- [Rud73] W. RUDIN – *Functional Analysis*, McGraw-Hill (1973).