ESAIM: Probability and Statistics November 1998, Vol. 2, 135–161

LARGE DEVIATIONS, CENTRAL LIMIT THEOREMS AND L^p CONVERGENCE FOR YOUNG MEASURES AND STOCHASTIC HOMOGENIZATIONS

JULIEN MICHEL AND DIDIER PIAU

Abstract. We study the stochastic homogenization processes considered by Baldi (1988) and by Facchinetti and Russo (1983). We precise the speed of convergence towards the homogenized state by proving the following results: (i) a large deviations principle holds for the Young measures; if the Young measures are evaluated on a given function, then (ii) the speed of convergence is bounded in every L^p norm by an explicit rate and (iii) central limit theorems hold. In dimension 1, we apply these results to the stochastic homogenization of random p-Laplacian operators for any $p > 1$.

1. Introduction

A non homogeneous material lies in a topological space Δ and its properties at a point x are described by the element $a(x)$ of a topological space Z. For instance, the function $a : \Delta \to Z$ may be the thermic conductivity of the material, its electric resistivity, or a deterministic function of these quantities. If the scale of the irregularities of the material is small, the function a is highly non regular. One can guess that the full knowledge of a, in other words a complete microscopic description of the material, is at the same time impossible to get and irrelevant if one is interested only in the macroscopic properties of the material. One way of avoiding this problem is to replace a by suitable random functions

$$a_\varepsilon(\omega, \cdot) : \Delta \to Z$$

where ω is the alea and ε is the typical scale of the irregularities of a_ε. When ε goes to zero, one expects that the behaviour of the random approximated material which is described by $a_\varepsilon(\omega, \cdot)$ converges, in a sense, to the behaviour of the actual material. This 'mean' description, often called stochastic homogenization, is intensely studied in the physical and mathematical literatures, see for example Kozlov (1980), Dal Maso and Modica (1986), Yurinskiĭ (1991) and the book of Jikov, Kozlov and Oleinik (1994). In the models that we study below, the convergence is well known or easy to establish and we precise the speed of this convergence to the homogenized state.

In fact, one often needs to study nonlinear quantities derived from the functions a_ε and their limit. A powerful approach is to study, instead of a_ε,

URL address of the journal: **http://www.emath.fr/ps/**

Received by the journal February 2, 1997. Revised April 9, 1998. Accepted for publication July 3, 1998.

the Young measure associated to a_ε. This is explained in Michel and Robert (1994) in the context of thermodynamical limits of infinite dimensional dynamical systems.

Fix a Borel probability measure dx on Δ. The Young measures ν that we are considering are defined, for any measurable $a : \Delta \to Z$, by

$$\nu(B) = \int_\Delta \mathbf{1}_B(x, a(x))\, dx$$

for any Borel set B of $\Delta \times Z$. The random Young measures ν_ε associated to the random functions a_ε often converge even if $a_\varepsilon(\omega, \cdot)$ doesn't converge in a usual sense.

We prove large deviations principles, central limit theorems and we bound the rate of the L^p convergence of ν_ε. Hence, we first study the asymptotic behaviour of

$$P(\nu_\varepsilon \in A),$$

for example when the measurable set A is asymptotically rare, that is when the limit ν_0 of ν_ε is not in A. Next, for regular enough functions $f : \Delta \times Z \to \mathbb{R}$, we show that the random variables $\varepsilon^{-1/2}(\nu_\varepsilon(f) - \nu_0(f))$ converge to a Gaussian law. Last, we bound the L^p norm of the random variables $(\nu_\varepsilon(f) - \nu_0(f))$ for every $p \geq 1$.

Random Young measures ν_ε contain all the information on the random functions a_ε. For instance, theorem 2.4 of Baldi (1988) is a large deviations principle for a_ε in model B (see the definition in section 2). Roughly speaking, Baldi's result states that, if a_ε is built using the law of a random variable α, then the exponential cost of a value $a(x) = t \in Z$ at the point $x \in \Delta$ is $j(t)$, where j is the Cramér transform of the law of α. Hence, $j(t)$ is exactly what it costs exponentially (*i.e.* the action of a large deviations principle) for i.i.d. random variables distributed like α to be such that their empirical mean is approximately equal to t instead of being approximately equal to the mean $E(\alpha)$ of α. Then, the exponential cost of a function $a : \Delta \to Z$ is the mean over Δ of $j \circ a$. We prove that this space-averaging of the cost (*i.e.* of the action of a large deviations principle) occurs for models B and FRp (see definitions in section 2) at the level of the Young measures. In other words, the total cost $i(\nu)$ of a measure ν is the mean over Δ of a local cost $c(\nu_x)$ which is measurable with respect to the value of ν at x. Furthermore, the function c is independent of x. For model B, we recover in section 3.3 the large deviations of a_ε (theorem 2.4 of Baldi (1988)) from the large deviations of ν_ε (our theorem 2.1). Baldi's result is obviously the source of our paper.

Space-averaging of the action of a large deviations principle is studied in Piau (1998) for a model of Poissonian obstacles in every dimension.

We mention that the large deviations principle for the Young measures of model B (our theorem 2.1) is proved by one of the authors and Robert in Michel and Robert (1994). We prove again this result for the following reasons: our assumption on the partitions of the topological space Δ (condition (P1) of section 2) is slightly weaker than in Michel and Robert (1994), our proof is slightly simpler (the only convexity property that we use is the concavity of the logarithm) and the application of theorem 2.1 to the large deviations of the Young measures $\nu_\varepsilon^{\mathrm{OP}}$ of the p-Laplacian with $p > 1$ (see the end of this section and the results of section 2.3) is not stated in the mentioned paper.

An important application of our results is the homogenization of random operators. In the following, Δ is often a domain of $\mathbb{R}^d$ of volume 1 and dx the Lebesgue measure. When $d = 1$, we proceed as follows. Assume that $\Delta = [0, 1]$, that α is a real valued random variable (that is, $Z = \mathbb{R}$) and that $v : \Delta \to \mathbb{R}$ is a given heat source of zero mean over Δ. Call u the temperature in the permanent regime of a material of thermic conductivity a, surrounded at $x = 0$ and $x = 1$ by an isolating material and submitted to the heat source v. Assume further that the heat flow q in the material follows a power law of the gradient u' of the temperature

$$q = -a\,|u'|^{r-2}\,u',$$

for $r > 1$. One recovers the usual Fourier law by setting $r = 2$. Then, u is the solution of

$$-(a(x)\,|u'(x)|^{r-2}\,u'(x))' = v(x), \quad x \in (0, 1),$$

with Neumann boundary conditions $u'(0) = u'(1) = 0$. The Young measure ν^{OP} associated to this r-Laplacian operator should describe at the same time the parameter a of the equation and the gradient u' of its solution. Hence, one sets

$$\nu^{\mathrm{OP}}(B) = \int_\Delta \mathbf{1}_B(x, a(x), u'(x))\,dx$$

for any Borel set B of $\Delta \times \mathbb{R}^2$. Of course, at least when a is nonnegative and bounded away from zero, u' is in fact

$$u'(x) = -|w(x)|^{q-1}\,w(x)\,a(x)^{-q}$$

where w is the primitive of v which is null at $x = 0$ and where $q = 1/(r-1)$. Hence, all the estimates on the Young measures ν, which takes only into account $(x, a(x))$, can be applied to the full Young measure ν^{OP} of the operator, which takes into account $(x, a(x), u'(x))$. We state the results for ν^{OP} in section 2.3.

In section 2, we recall the stochastic homogenization processes due to Baldi (1988) and to Facchinetti and Russo (1983) and we state the results on ν and on ν^{OP}. Notice that we study the model introduced in dimension 1 by Baldi, in a general setting. Finally, we give the plan of the rest of the paper.

2. Results

Before introducing the two models, we define some conventions related to the statements of large deviation principles. The relation

$$P(\nu_n \in A) \sim \exp(-n\,i(A))$$

is a concise form of the following assertion. The sign $\sim$ stands for the two bounds of “à la Varadhan” large deviations principles: for any open set G and any closed set F, we ask that

$$\liminf_{n\to\infty} n^{-1} \log P(\nu_n \in G) \geq -i(G),$$
$$\limsup_{n\to\infty} n^{-1} \log P(\nu_n \in F) \leq -i(F),$$

where, for any set A, $i(A)$ stands for

$$i(A) = \inf\{i(\nu)\,;\,\nu \in A\}.$$

The action i takes its value in $[0, +\infty]$ and its level lines $i^{-1}([0,t])$ are compact. (This implies that i is lower semi continuous.)

All the large deviations principles which are stated below for Young measures are valid in the space of bounded Borel measures with the topology of narrow convergence. All the large deviations principles for volumes of sets are valid in $\mathbb{R}$ with the usual topology.

2.1. The model B

In model B, studied in dimension 1 by Baldi (1988) and in a general setting by Michel and Robert (1994), the scale of the irregularities, for a given scale of discretization, is constant through the material and the value of the parameter on each cell delimited by the irregularities is random.

Set a locally compact, separable metrisable space Δ and a Borel probability measure dx on Δ (or a domain Δ of $\mathbb{R}^d$ of volume $|\Delta| = 1$ and dx the Lebesgue measure). Choose a sequence of measurable partitions $(\Delta_n)_n$ of the domain Δ with $\Delta_n = \{\Delta_i^n\}_i$ and consider the following properties.

Call Δ_n^* the union of the cells Δ_i^n of volume exactly n^{-1}, v_n^* the volume of $\Delta \setminus \Delta_n^*$, δ_n^* the maximum of the diameters of the cells in Δ_n^* and δ_n the maximum of all the diameters of the cells of the partition Δ_n.

Property (P1): v_n^* and δ_n^* converge to zero as n goes to infinity.

Property (P2): for given nonnegative constants c and d, the volume of any cell of Δ_n is at most n^{-1}, its diameter is at most $c\,n^{-1/d}$ and the total number of cells $N(n)$ is less than $c\,n$.

Property (P3): the volume of any cell of Δ_n is at most n^{-1}, and δ_n and v_n^* converge to zero.

Any of these properties implies that dx is diffuse. When d is an integer, property (P2) states that Δ behaves like a domain of $\mathbb{R}^d$. If Δ is a domain of $\mathbb{R}^d$ and if the area of its boundary is finite, sequences of partitions satisfying (P1-2-3) exist: take for the cells of Δ_n the non void intersections of Δ with the cubes of the grid $n^{-1/d}\,\mathbb{Z}^d$. Condition (P1) follows the idea of Michel and Robert (1994) but to check (P1) is a little easier than to check their equipartition condition.

Fix once for all a random variable α with values in a Polish space Z (separable metrisable complete topological space) endowed with its Borel σ-field. Let $(\alpha_i)_i$ be i.i.d. copies of α. Define $a_n(x) = \alpha_i$ if $x \in \Delta_i^n$, that is

$$a_n(\omega, x) = \sum_{i \geq 1} \alpha_i(\omega)\,\mathbf{1}_{\Delta_i^n}(x),$$

and let ν_n be the Young measure describing a_n. Define a Young measure ν_∞ by

$$\nu_\infty(B) = \int_\Delta E(\mathbf{1}_B(x, \alpha))\,dx.$$

For any measures μ and ν, the relative entropy $\mathcal{K}(\mu|\nu)$ (or Kullback's information) of μ with respect to ν is

$$\mathcal{K}(\mu|\nu) = \int \log(d\mu/d\nu)\,d\nu$$

if the Radon-Nikodym derivative $d\mu/d\nu$ exists, and $\mathcal{K}(\mu|\nu) = +\infty$ otherwise.

Theorem 2.1 is essentially due to Michel and Robert (1994) and it is a generalization of theorem 2.4 in Baldi (1988). The general definition of a Young measure is in section 3.1.

THEOREM 2.1. *Assume that (P1) holds. The Young measures ν_n of model B converge a.s. to ν_∞. They satisfy a large deviation principle*

$$P(\nu_n \in A) \sim \exp(-n\, i(A))$$

of rate $i(\nu) = \mathcal{K}(\nu|\nu_\infty)$ if ν is a Young measure absolutely continuous with respect to ν_∞ and $i(\nu) = +\infty$ otherwise.
Notice that $i(\nu) = +\infty$ if ν is not a Young measure, even if ν is absolutely continuous with respect to ν_∞. Call P_α the law of α. Then, if ν is a Young measure of disintegration $(\nu_x)_x$,

$$i(\nu) = \int_\Delta \mathcal{K}(\nu_x|P_\alpha)\, dx.$$

THEOREM 2.2.
i. Assume that (P2) holds. Let f be any bounded Hölder continuous function of exponent β. Set $\gamma = \min(\beta/d, 1/2)$. For every real $p \geq 0$, there exists a constant $c_p(f) = c_p$ such that, for any $n \geq 1$, one has

$$E(|\nu_n(f) - \nu_\infty(f)|^p) \leq c_p\, n^{-\gamma\, p}.$$

ii. Assume now that (P3) holds. For any bounded continuous function f,

$$n^{1/2}(\nu_n(f) - \nu_\infty(f))$$

converges in law to a centered Gaussian random variable of variance

$$\int_\Delta \mathrm{Var}\,(f(x,\alpha))\, dx, \quad \mathrm{Var}\,(h(\alpha)) = E([h(\alpha) - E(h(\alpha))]^2).$$

2.2. THE MODEL FR

In model FR, due to Facchinetti and Russo (1983), Δ is a domain of $\mathbb{R}$, dx is the Lebesgue measure and the size of the cells delimited by the irregularities is random. In fact, all the randomness of the material is concentrated there and not in the values of the parameter $a(x)$. Fix for example $\Delta = [0,1]$. Choose once for all two values of $a(x)$, say a^+ and a^-, and define

$$a = a^+\,\mathbf{1}_{\Delta^+} + a^-\,\mathbf{1}_{\Delta^-} \tag{2.1}$$

where $\{\Delta^+, \Delta^-\}$ is a random partition of Δ. More precisely, one sets

$$a(\omega, x) = a^+\,\mathbf{1}_{\Delta^+(\omega)}(x) + a^-\,\mathbf{1}_{\Delta^-(\omega)}(x).$$

Choose n points at random in Δ with the uniform law on Δ^n. The value of a is alternatively a^+ and a^- on each of the $(n+1)$ intervals of Δ delimited by these points. More precisely, if $x_1 < x_2 < \cdots < x_n$ are the points chosen at random, set $x_0 = 0$ and $x_i = 1$ for $i \geq n+1$. Then,

$$\Delta_n^+ = \bigcup_{i\geq 0} [x_{2i}, x_{2i+1}[, \quad \Delta_n^- = \Delta \setminus \Delta_n^+.$$

This is the model of Facchinetti and Russo. We modify their construction as follows. Introduce a Poisson point process on Δ of constant intensity λ and replace the n points chosen at random by the points of this Poisson process, whose number is random. Leave the rest of the construction as it was. We call this model, Poissonian model of Facchinetti and Russo (FRp). Of course, the set of the points of a Poisson process, conditioned by its cardinal, follows the law of the uniform sample used in FR. We write $\Delta^{\pm}_{\lambda}$ for the partition of Δ when the intensity of the Poisson process is λ, a_λ for the function defined by (2.1) with this partition and ν_λ for the corresponding Young measure. The subscript n instead of λ refers to FR instead of FRp.

Some notation is needed before we state our results. For any measurable function $p : \Delta \to [0,1]$, ν^p is the Young measure

$$\nu^p(B) = \int_\Delta [p(x)\,\mathbf{1}_B(x,a^+) + (1-p(x))\,\mathbf{1}_B(x,a^-))]\,dx.$$

Denote $\nu^+ = \nu^1$ and $\nu^- = \nu^0$. The two Young measures $\nu^\pm$ describe a homogeneous material of constant parameter $a^\pm$. The Young measure ν^p describes a material, fictitious if p is not always 0 or 1, where each point x is a mixture of a^+ and a^- in proportions $p(x)$ and $1-p(x)$.

The convex set of the Young measures ν^p for all measurable $p : \Delta \to [0,1]$ is the closure for the topology of the narrow convergence of the set of the Young measures which describe a given real material, that is of the measures ν^p where p takes the values 0 and 1 only. Hence, this convex set contains the domain of the action of any large deviations principle for ν_n or for ν_λ.

Set $c(y) = 1 - \sqrt{4y(1-y)}$ if $y \in [0,1]$ and $c(y) = +\infty$ otherwise.

Theorem 2.3. (Model FRp) *The Lebesgue measure of any of the two components of the random partition, for instance Δ^+_λ, converges a.s. to $\frac{1}{2}$ and satisfies the large deviations principle*

$$P(|\Delta^+_\lambda| \in A) \sim \exp(-\lambda\, c(A)).$$

The Young measures ν_λ converge a.s. to $\nu^{1/2} = \frac{1}{2}(\nu^+ + \nu^-)$ and they satisfy the large deviations principle

$$P(\nu_\lambda \in A) \sim \exp(-\lambda\, i(A)).$$

The action i is defined by

$$i(\nu^p) = \int_\Delta c(p(x))\,dx$$

for any measurable $p : \Delta \to [0,1]$ and $i(\nu) = +\infty$ otherwise.

Theorem 2.4. (Model FRp) *Let f be a bounded Hölder continuous function of exponent β. For every integer $p \geq 0$, there exists a constant $c_p(f) = c_p$ such that, for any $\lambda > 0$, one has*

$$E(|\nu_\lambda(f) - \nu^{1/2}(f)|^p) \leq c_p\,\lambda^{-p\beta/(p+2\beta)}.$$

Let f be a piecewise constant function. Then, $\lambda^{1/2}(\nu_\lambda(f) - \nu^{1/2}(f))$ converges in law to a centered Gaussian random variable of variance

$$\int_\Delta g^2(x)\,dx \quad \text{with} \quad g(x) = \tfrac{1}{2}(f(x,a^+) - f(x,a^-)).$$

Here is the non Poissonian version of theorems 2.3 and 2.4.

THEOREM 2.5. (MODEL FR) *The Lebesgue measure of any of the two components of the random partition, for instance* Δ_n^+*, converges a.s. to* $\frac{1}{2}$ *and satisfies the large deviations principle*

$$P(|\Delta_n^+| \in A) \sim \exp(-n\, c_{\rm DET}(A)), \quad c_{\rm DET} = -\log(1-c).$$

The Young measures ν_n *converge a.s. to* $\nu^{1/2} = \frac{1}{2}(\nu^+ + \nu^-)$ *and they satisfy the large deviations principle*

$$P(\nu_n \in A) \sim \exp(-n\, i_{\rm DET}(A)).$$

The action $i_{\rm DET}$ *is defined by*

$$i_{\rm DET}(\nu^p) = -\log \int_\Delta [1 - c(p(x))]\, dx = -\log(1 - i(\nu^p))$$

for any measurable $p : \Delta \to [0,1]$ *and* $i_{\rm DET}(\nu) = +\infty$ *otherwise. Notice that* $i_{\rm DET}(\nu) = +\infty$ *if* ν *describes a given deterministic material, that is if* $\nu = \nu^p$ *for* p *taking only the values* 0 *and* 1*.*
The L^p *bound of theorem 2.4 holds for any integer* $p \geq 1$ *if* n *replaces* λ *everywhere and if the bound* $c_p\, n^{-\beta/(2+\beta)}$ *replaces the bound* $c_p\, \lambda^{-p\beta/(p+2\beta)}$*.*

We have not proved the analog of the central limit theorem of theorem 2.4 for the model FR because of technical difficulties, although it presumably holds.

Theorems 2.3, 2.4 and 2.5 are valid with small modifications if one replaces the two real numbers $a^\pm$ by two functions $a^\pm : \Delta \to \mathbb{R}$. In the L^p estimates, $f(\cdot, a^\pm(\cdot))$ (instead of f) must be β Hölder. In the central limit theorem, $f(\cdot, a^\pm(\cdot))$ (instead of f) must be piecewise constant. The modifications of the large deviations result are stated in section 5.5.

2.3. HOMOGENIZATION OF OPERATORS

We resume the situation of the end of Section 1 and apply the previous results to the Young measure $\nu^{\rm OP}$ of the r-Laplacian. Assume that $a \geq a_*$ for a given real $a_* > 0$. Recall that the dimension $d = 1$, that $\Delta = [0,1]$ and that u' is

$$u'(x) = -|w(x)|^{q-1}\, w(x)\, a(x)^{-q}$$

with $q = 1/(r-1)$. Assume that w is bounded and write $h = -|w|^{q-1}\, w$. For a given function a, $\nu^{\rm OP}$ is defined by

$$\nu^{\rm OP}(B) = \int_\Delta \mathbf{1}_B(x, a(x), u'(x))\, dx = \int_\Delta \mathbf{1}_B(x, a(x), h(x)\, a(x)^{-q})\, dx.$$

Conditions (P1-2-3) hold for $\Delta = [0,1]$ and the canonical regular partition given by $\Delta_i^n = [(i-1)/n, i/n)$.

PROPOSITION 2.6. (MODEL B) *Assume that* $\alpha \geq a_* > 0$ *a.s. and define*

$$\nu_\infty^{\rm OP}(B) = \int_\Delta E[\mathbf{1}_B(x, \alpha, h(x)\, \alpha^{-q})]\, dx.$$

Then, $(\nu_n^{\rm OP})_n$ *satisfies a large deviations principle*

$$P(\nu_n^{\rm OP} \in A) \sim \exp(-n\, i^{\rm OP}(A)),$$

where the action is $i^{\rm OP}(\nu) = \mathcal{K}(\nu | \nu_\infty^{\rm OP})$ *if* ν *is a Young measure and* $i^{\rm OP}(\nu) = +\infty$ *otherwise.*

REMARK 2.7. Assume that $\nu^{\rm OP}$ and ν are defined by the same function a. Then, one has

$$\mathcal{K}(\nu^{\rm OP}|\nu^{\rm OP}_\infty) = \mathcal{K}(\nu|\nu_\infty).$$

PROPOSITION 2.8. (MODEL B) *Let f be Hölder continuous of exponent β. Recall that $q = 1/(r-1)$ and set $\gamma = \min(\beta, \beta q, \frac{1}{2})$. For every real $p \geq 0$, there exists a constant $c_p(f) = c_p$ such that, for any $n \geq 1$, one has*

$$E(|\nu^{\rm OP}_n(f) - \nu^{\rm OP}_\infty(f)|^p) \leq c_p\, n^{-\gamma p}.$$

For any bounded continuous function f, $n^{1/2}(\nu^{\rm OP}_n(f) - \nu^{\rm OP}_\infty(f))$ converges in law to a centered Gaussian random variable of variance

$$\int_\Delta \mathrm{Var}\,(f(x, \alpha, h(x)\,\alpha^{-q}))\, dx.$$

For the model FRp, let $a^\pm$ be two a.e. continuous functions with $a^\pm \geq a_*$ and let Δ_e be the set of the points $x \in \Delta$ where $a^+(x) = a^-(x)$. Call admissible any measurable function $p : \Delta \to [0,1]$ such that $p = \frac{1}{2}$ on Δ_e (see Section 5.5). For any admissible p, define

$$\begin{aligned}\nu^p_{\rm OP}(B) \quad &= \quad \int_\Delta [p(x)\,\mathbf{1}_B(x, a^+(x), h(x)\,a^+(x)^{-q}) \\ &\qquad + (1-p(x))\,\mathbf{1}_B(x, a^-(x), h(x)\,a^-(x)^{-q})]\, dx,\end{aligned}$$

and denote by $\nu^+_{\rm OP} = \nu^1_{\rm OP}$ and $\nu^-_{\rm OP} = \nu^0_{\rm OP}$.

PROPOSITION 2.9. (MODELS FRP AND FR) *The Young measures $\nu^{\rm OP}_\lambda$ of the model FRp converge a.s. to $\nu^{1/2}_{\rm OP} = \frac{1}{2}(\nu^+_{\rm OP} + \nu^-_{\rm OP})$ and they satisfy the large deviations principle*

$$P(\nu^{\rm OP}_\lambda \in A) \sim \exp(-\lambda\, i^{\rm OP}(A)),$$

where the action $i^{\rm OP}$ is defined by $i^{\rm OP}(\nu^p_{\rm OP}) = i(\nu^p)$ for any admissible p and $i^{\rm OP}(\nu) = +\infty$ otherwise. The same result holds for the Young measures $\nu^{\rm OP}_n$ of the model FR with an action $i^{\rm OP}_{\rm DET}$ defined by

$$i^{\rm OP}_{\rm DET}(\nu^p_{\rm OP}) = -\log(1 - i^{\rm OP}(\nu^p_{\rm OP})) = i_{\rm DET}(\nu^p)$$

for any admissible $p : \Delta \to [0,1]$ and $i^{\rm OP}_{\rm DET}(\nu) = +\infty$ otherwise.

The central limit theorem and the L^p estimates of theorem 2.4 can also be adapted to this setting.

2.4. APPLICATIONS OF THE CONVERGENCE OF THE YOUNG MEASURES

The convergence of $\nu^{\rm OP}$ implies the convergence of a and u' as follows. Denote by $\widetilde{\nu}$ the limit of $\nu^{\rm OP}_n$ or of $\nu^{\rm OP}_\lambda$ in one of the models. Hence, $\widetilde{\nu} = \nu^{\rm OP}_\infty$ in model B and $\widetilde{\nu} = \nu^{1/2}_{\rm OP}$ in models FR and FRp. Call $\widetilde{a}$ the function such that $\widetilde{a}(x)^{-q} = E(\alpha^{-q})$ for model B, and $\widetilde{a}(x)^{-q} = \frac{1}{2}(a^+(x)^{-q} + a^-(x)^{-q})$ for models FR and FRp. Except for the degenerate case where α is a.s. constant, or where $a^+ = a^-$ a.s., $\widetilde{\nu}$ is not the Young measure describing the operator of coefficient $\widetilde{a}$.

If one stays at the level of a, consider model B as in Baldi (1988), that is with $\Delta = [0,1]$ and $Z = \mathbb{R}$. Assume that α is bounded and call m the Cramér transform of the law of α, that is

$$m(t) = \sup\{st - \log E(\exp(s\,\alpha))\,;\, s \in \mathbb{R}\}.$$

Baldi proved that the law of a_n converges in L^∞ for the $\sigma(L^\infty, L^1)$ topology and satisfies a large deviations principle

$$P(a_n \in A) \sim \exp(-n\, i_B(A)), \qquad i_B(f) = \int_\Delta m(f(x))\, dx.$$

This result is motivated by problems of Γ-convergence of some operators, see Baldi (1988), Dal Maso and Modica (1986) and Dal Maso (1993). Going back to the Young measures and to any of the models B, FR and FRp, our large deviations results imply the following proposition. See also Section 3.3 where we deduce the large deviations principle due to Baldi in $\Delta = [0, 1]$ for the functions a_n in model B, from our Theorem 2.1, and Section 5.4 where we deduce a large deviations principle for the functions a_λ in model FR from our Theorem 2.3.

Proposition 2.10.
i. A.s., ν_n^{OP} converges narrowly to $\widetilde{\nu}$.
Assume that $\alpha \geq a_* > 0$ in model B and that $a^\pm > 0$ in model FR.
ii. A.s., a_n^{-q} converges to the deterministic function $\widetilde{a}^{-q}$ in the $\sigma(L^\infty, L^1)$ topology.
iii. The solution u_n converges a.s. weakly to the solution $\widetilde{u}$ of the deterministic problem with coefficient $\widetilde{a}$, that is

$$-(\widetilde{a}(x)\, |\widetilde{u}'(x)|^{r-2}\, \widetilde{u}'(x))' = v.$$

Proof. *i* and *ii* are direct. For *iii*, choose $f(x, y, z) = z\, h(x)$ with h integrable and introduce

$$A = \{\nu\, ;\, |\nu(f) - \widetilde{\nu}(f)| \geq t\}.$$

If $i^{\mathrm{OP}}(A) > 0$, $P(\nu_n^{\mathrm{OP}} \in A)$ is summable by the large deviations principle. By Borel-Cantelli's lemma, for n large enough, $\nu_n^{\mathrm{OP}} \notin A$, that is $|\nu_n^{\mathrm{OP}}(f) - \widetilde{\nu}(f)| < t$. We proved that $\nu_n^{\mathrm{OP}}(f) - \widetilde{\nu}(f)$ goes to zero. Furthermore,

$$\nu_n^{\mathrm{OP}}(f) - \widetilde{\nu}(f) = \int_\Delta h(x)\, (u_n'(x) - \widetilde{u}'(x))\, dx,$$

hence u_n' converges a.s. weakly to $\widetilde{u}'$. The last step is to prove that $i^{\mathrm{OP}}(A) > 0$. In model B, use the general relation (Dembo and Zeitouni (1992), exercise 6.2.17):

$$\mathcal{K}(\nu|\mu) \geq \tfrac{1}{2}(\mu(f) - \nu(f))^2$$

for any measures μ and ν and any measurable f bounded by 1. In model FR or FRp, by Cauchy-Schwarz inequality, $\nu_{\mathrm{OP}}^p \in A$ implies that the mean of $(p - \frac{1}{2})^2$ is at least $c\, t^2$, hence that $(p - \frac{1}{2})^2 \geq \varepsilon$ on a set of measure at least ε and this yields an explicit positive lower bound of $i^{\mathrm{OP}}(\nu_{\mathrm{OP}}^p)$. □

2.5. Plan

The rest of the paper is organized as follows. In Section 3, we prove the large deviations principle for model B (Theorem 2.1), in Section 4 the L^p estimate and central limit theorem for model B (Theorem 2.2), in Section 5 the large deviations principle for model FRp (Theorem 2.3) and in Section 6, the L^p estimate and central limit theorem for model FRp. A part of Section 3 is used in Section 5. Theorem 2.5 is proved in Section 6.4 for the L^p estimates. For the large deviations principle, technical difficulties arise and we give the proof in the Appendix.

3. Proof of theorem 2.1

3.1. A general reduction

We first collect general facts about large deviations in topological vector spaces and we apply them to the proof of Theorem 2.1. We will use the results of this section in the proof of Theorem 2.3 (Section 5.2).

Let E be a locally convex topological vector space, E^* be its topological dual and μ_n be probability measures on E. Assume that:
(H1) the family $(\mu_n)_n$ is exponentially tight;
(H2) for any $f \in E^*$, $n^{-1} \log \mu_n(\exp(nf))$ converges to a limit $l(f)$;
(H3) the function l is finite valued and Gateaux differentiable on E^*.

Then (see corollary 4.6.14 of Dembo and Zeitouni (1992)), $(\mu_n)_n$ satisfies a large deviations principle

$$\mu_n(A) \sim \exp(-n\, i(A))$$

where the action i is the Legendre transform of l, that is

$$i(\nu) = \sup\{\langle f, \nu\rangle - l(f)\,;\, f \in E^*\}.$$

We choose for E the space of bounded measures on $\Delta \times Z$ with the topology of the vague convergence and for μ_n the law of ν_n. Denote by F the space of bounded measures on $\Delta \times Z$ with the weak topology (*i.e.* the topology of the narrow convergence): our large deviations principles are stated in F. We show that the sequence μ_n satisfies (H1) in E and how to deduce a large deviations principle in F from a principle in E. This part of the proof applies to Theorem 2.3.

First, introduce the subset Y of E of the Young measures of base the measure dx on Δ. Recall that $\nu \in Y$ iff there exists a measurable collection $(\nu_x)_x$ of probability measures such that

$$\nu(f) = \int_\Delta \nu_x(f(x, \cdot))\, dx$$

for all bounded measurable f. Such a collection $(\nu_x)_x$ is called a disintegration of ν. Since Z is a Polish space, regular conditional distributions of the Z coordinate always exist (see for example Breiman (1968)). Hence, $\nu \in Y$ iff the first marginal of ν is the measure dx, that is iff, for any Borel subset A of Δ,

$$\nu(A \times Z) = \int_\Delta \mathbf{1}_A(x)\, dx = |A|. \tag{3.1}$$

If Z is compact, Y is closed in E. If Z is not compact, the closure of Y in E is the set of the measures $\nu \in E$ such that $\nu(A \times Z) \leq |A|$ for any Borel subset of Δ. In both cases, the closure of Y in E is compact in E and the complementary set has zero measure with respect to each μ_n, hence (H1) holds.

Assume now that $(\mu_n)_n$ satisfies a large deviations principle in E and that, for all n, $\mu_n(Y^c) = 0$. For any open subset G of F, $\mu_n(G) = \mu_n(G \cap Y)$ and $G \cap Y$ is open in Y for the narrow topology, hence for the vague one since both topologies coincide on a set of probability distributions. There exists G' open in E such that $G \cap Y = G' \cap Y$, so that $\mu_n(G) = \mu_n(G')$.

Now,

$$\liminf_{n\to\infty} n^{-1} \log \mu_n(G') \geq -i(G') = -i(G' \cap Y) = -i(G)$$

because $i = +\infty$ on Y^c, hence the lower bound holds for any open set in F. The same method gives the upper bound for the closed sets in F.

By the preceding, we only have to prove that the laws of ν_n satisfy (H2) and (H3) for the vague topology: we can identify the dual of E with the space $C_K(\Delta \times Z) = C_K$ of continuous functions with a compact support through the duality relation $\langle f, \nu \rangle = \nu(f)$.

3.2. (H2) and (H3) hold in model B

(P1) holds. For $f \in C_K$, we estimate

$$\mu_n(\exp(nf)) = E(\exp(n\,\nu_n(f))).$$

First, since f is uniformly continuous and since the diameter of the cells of Δ_n^* goes uniformly to zero, there exists a function f_n which is constant in the first coordinate on each cell of Δ_n^* and such that the uniform norm of $f - f_n$ on Δ_n^* goes to zero with n. Call this uniform norm ε_n and $f_i^n(\cdot) = f_n(x, \cdot)$ for $x \in \Delta_i^n$ with $\Delta_i^n \subset \Delta_n^*$. Set $f_n(x, \cdot) = 0$ if $x \in \Delta \setminus \Delta_n^*$.
Replacing Δ by Δ_n^* and f by f_n in $\mu_n(\exp(nf))$ produces an error of at most

$$\exp O(n\,(\varepsilon_n + v_n^*)).$$

We write the mean over Δ_n^* defining $\nu_n(f_n)$ as the sum of the means over the cells Δ_i^n and we use the independence of the random variables α_i, getting

$$\begin{aligned}\mu_n(\exp n\,f_n) &= E\left[\exp\left(n\int_{\Delta_n^*} f_n(x, a_n(x))\,dx\right)\right] \\ &= \prod_i E\left[\exp\left(n\,|\Delta_i^n|\,f_i^n(\alpha)\right)\right]\end{aligned}$$

where the product is restricted to the cells of Δ_n^*. Taking logarithms, one gets:

$$\log \mu_n(\exp nf) = O(n\,(\varepsilon_n + v_n^*)) + \sum_i \log E(\exp f_i^n(\alpha)),$$

where the sum is restricted to the cells of Δ_n^*. The sum over i is n times the mean of $\log E(\exp(f_n(\cdot, \alpha))$ over Δ_n^*. Replacing f_n by f and Δ_n^* by Δ in this mean causes another error of $n\,(\varepsilon_n + v_n^*)$, hence the errors disappear when divided by n. Finally, (H1) holds with

$$l(f) = \int_\Delta \log E(\exp f(x, \alpha))\,dx.$$

The functional l is finite everywhere, convex and Gateaux differentiable by the next lemma, so that (H3) holds. Lemma 3.1 also provides explicitly the action i of theorem 2.1 by computing the Legendre transform of a more general functional (see also Michel and Robert (1994)).

Lemma 3.1. *Let Δ be a locally compact, separable metrisable topological space and dx a probability measure on Δ such that condition (P1) holds. Let Z be a Polish space and π be a Young measure on $\Delta \times Z$ of base dx and* *disintegration* $(\pi_x)_x$. *For any* $x \in \Delta$, *let* α_x *be a random variable with value in* Z *of law* π_x. *Let* l *be defined on* $C_K(\Delta \times Z)$ *by*

$$l(f) = \int_\Delta \log E(\exp(f(x, \alpha_x)))\, dx.$$

Then, l *is finite, convex and differentiable on* $C_K(\Delta \times Z)$. *Furthermore, the Legendre transform* i *of* l *is* $i(\nu) = \mathcal{K}(\nu|\pi)$, *the relative entropy of* ν *with respect to* π, *if* ν *is a Young measure, and* $i(\nu) = +\infty$ *otherwise.*
When ν *is a Young measure of disintegration* $(\nu_x)_x$, *one has*

$$i(\nu) = \int_\Delta \mathcal{K}(\nu_x|\pi_x)\, dx.$$

Proof. By definition, the Legendre transform i of l is

$$i(\nu) = \sup\{\nu(f) - l(f)\,;\, f \in C_K\}.$$

The finiteness of l on C_K is direct and l is convex as a mean of convex functions. The differential $dl(f)$ of l at $f \in C_K$ can be computed explicitly. For any $g \in C_K$, one gets

$$\langle dl(f), g\rangle = \int_\Delta \frac{E(g(x, \alpha_x)\exp f(x, \alpha_x))}{E(\exp f(x, \alpha_x))}\, dx.$$

We compute i in several steps.
Step 1. If ν is not a positive measure, let $f \leq 0$ such that $\nu(f) > 0$. Then $l(f) \leq 0$, hence $i(\nu)$ is at least $\nu(tf) = t\,\nu(f)$ for any t. This proves that $i(\nu)$ is infinite.
Step 2. Let ν be a positive measure such that $i(\nu)$ is finite; ν must be finite. For $h \in C_K(\Delta)$ and $k \in C_K(Z)$, $h \otimes k \in C_K$. Hence,

$$i(\nu) \geq t\,\nu(h \otimes k) - l(t(h \otimes k))$$

for any t. Fix t and choose for k a sequence increasing to $\mathbf{1}_Z$. Since ν is finite and h and k are bounded, the two terms of the difference converge by Lebesgue monotone convergence theorem to finite limits, hence

$$i(\nu) \geq \nu(t(h \otimes \mathbf{1})) - l(t(h \otimes \mathbf{1})) = t\left(\nu(h \otimes \mathbf{1}) - \int_\Delta h(x)\, dx\right)$$

for all real t. This shows that the first projection of ν is the Lebesgue measure. By the characterization of equation (3.1), this implies that ν is a Young measure.
Step 3. Let ν be any measure and $f \in C_K$. By convexity of the exponential,

$$l(f) \leq \log\left(\int_\Delta \pi_x(\exp f(x, \cdot))\, dx\right) = \log \pi(\exp f),$$

hence $i(\nu) \geq \mathcal{K}(\nu|\pi)$ for any measure ν, from the variational characterization of the entropy $\mathcal{K}$ that we recall in the next step of this proof.
Step 4. Let ν be a Young measure of disintegration $(\nu_x)_x$. Recall that $\mathcal{K}(\nu_x|\pi_x)$ can be characterized as

$$\mathcal{K}(\nu_x|\pi_x) = \sup_g\,[\nu_x(g) - \log \pi_x(\exp g)]$$

for all $g \in C_K(Z)$, see for example Dembo and Zeitouni (1992), lemma 6.2.13. For any $f \in C_K$, each $f(x,\cdot) \in C_K(Z)$, hence

$$\begin{aligned} \nu(f) - l(f) &= \int_\Delta [\nu_x(f(x,\cdot)) - \log \pi_x(\exp f(x,\cdot))]\, dx \\ &\leq \int_\Delta \mathcal{K}(\nu_x|\pi_x)\, dx = \mathcal{K}(\nu|\pi), \end{aligned}$$

from the definition of the relative entropy using the Radon-Nikodym derivatives. This shows that $i(\nu) \leq \mathcal{K}(\nu|\pi)$ for any Young measure ν. □

3.3. Large deviations of a_n

Assume for simplicity that $Z = \mathbb{R}$. Assume, more importantly, that α is bounded. Notice that $a_n = f$ if and only if ν_n is a Young measure ν such that, for all $g : \Delta \to \mathbb{R}$, $\nu(\bar{g})$ is the mean over Δ of fg, where $\bar{g}(x,y) = y\, g(x)$. This amounts to say that a disintegration $(\nu_x)_x$ of ν satisfies $\nu_x(y) = f(x)$ for all $x \in \Delta$. Because α is bounded, the function y can be replaced by a bounded continuous function. This yields a contraction principle as follows.

Since $i(\nu) = +\infty$ for $\nu \notin Y$, a large deviations principle holds on Y. Let T be a continuous bounded function such that $T(y) = y$ on the support of α. Define $\Phi : Y \to L^\infty(dx)$ by

$$\Phi(\nu)(x) = E_\nu(T(y)|x) = \nu_x(T).$$

Then, Φ is continuous on Y for the $\sigma(L^\infty(dx), L^1)$ topology. Hence, the contraction principle is valid and a large deviations principle for a_n holds as

$$P(a_n \in A) \sim \exp(-n\, i_B(A))$$

where the action i_B is defined as follows: $i_B(f)$ is the infimum of $i(\nu)$ over all the Young measures ν of disintegration $(\nu_x)_x$ such that $\nu_x(y) = f(x)$ for all $x \in \Delta$. Moreover, on the set of the Young measures, i is the relative entropy with respect to ν_∞ and the disintegration of ν_∞ is $\nu_x^\infty = P_\alpha$ for all $x \in \Delta$, where P_α is the law of α. Hence,

$$i_B(f) = \int_\Delta j(f(x))\, dx$$

if $j(t)$ is defined by

$$j(t) = \inf\{\mathcal{K}(\mu|P_\alpha)\,;\, \mu(y) = t\}.$$

Finally, notice that j is also the Cramér transform of P_α, that is

$$j(t) = \sup\{xt - \log P_\alpha(\exp(xy))\,;\, x \in \mathbb{R}\}.$$

At least when the support of P_α is bounded, this is a consequence of Sanov's theorem and of the contraction principle, as follows. Since α is bounded, there exists $f \in C_K$ such that $\mu(y) = \mu(f)$ for any μ absolutely continuous with respect to P_α. Hence, $\mu \mapsto \mu(y)$ is continuous on the domain of $\mathcal{K}(\cdot|P_\alpha)$ and the contraction principle is valid. This proves that the action deduced by contraction from Sanov's theorem is also the action of Cramér's theorem for i.i.d. random variables, that is the Cramér transform of P_α.

4. Proof of theorem 2.2

We first reduce the problem and set some notations, useful for the L^p estimates as well as for the central limit theorem. Assume without loss of generality that $|f|$ is uniformly bounded by 1 and that $f(x, \alpha)$ is centered for every $x \in \Delta$. Approximate f by a function f_n which is constant on each Δ_i^n and such that $f_n(x, \alpha)$ is centered for every $x \in \Delta$. Since the diameter of each cell is at most δ_n going to zero and since f is uniformly continuous, one can ask that the uniform norm ε_n of $f - f_n$ converges to zero. The uniform norm of f_n is at most a constant. If $\delta_n \leq c\, n^{-1/d}$ and f is β Hölder, then ε_n is $O(n^{-\beta/d})$.

Write $f_i^n(\cdot) = f_n(x, \cdot)$ when $x \in \Delta_i^n$. For $n \geq 1$, $\nu_n(f_n)$ is the sum of the independent random variables

$$|\Delta_i^n|\, f_i^n(\alpha_i).$$

4.1. L^p estimates

Assumption (P2) holds and f is β Hölder. From the preceding, one has

$$E(|\nu_n(f)|^p) \leq c\, n^{-p\beta/d} + E(|\nu_n(f_n)|^p) = cn^{-p\beta/d} + E\left(\left|\sum_{i\geq 1} |\Delta_i^n|\, f_i^n(\alpha_i)\right|^p\right)$$

We can assume that p is an even integer since the desired bound for p implies the desired bound for any $p' \leq p$ by Hölder convexity inequality. Then, the p-th power can be written without the absolute value sign. In the multinomial expansion of the p-th power of the sum, the expectations can be factorized because the α_i are independent, and all the first moments disappear because each $f_i^n(\alpha)$ is centered. We bound all the other terms by the p-th power of the infinite norm of f_n. Since the volume of a cell is at most n^{-1}, one gets

$$E\left(\left(\sum_{i\geq 1} |\Delta_i^n|\, f_i^n(\alpha_i)\right)^p\right) \leq c\, n^{-p}\, A_n,$$

where A_n is the number of terms without any first moment. Calling $N(n)$ the number of cells in Δ_n, A_n is the number of integer-valued multi-indices r of size $N(n)$ and of sum p such that $r_i = 0$ or $r_i \geq 2$ for all i. There is at most $c\, N(n)^{p/2}$ such multi-indices and $N(n) \leq c\, n$, hence

$$E(|\nu_n(f)|^p) \leq c\, (n^{-p\beta/d} + n^{-p/2}).$$

4.2. Central limit theorem

Assumption (P3) is satisfied. Since f is bounded, the random variables

$$X_{i,n} = n^{1/2} \int_{\Delta_i^n} f(x, \alpha_i)\, dx$$

are independent and uniformly $O(n^{-1/2})$. It follows that, as soon as the variance of $n^{1/2}\, \nu_n(f)$ converges to a strictly positive constant, $(X_{i,n})_i$ satisfies Lindeberg condition and Lindeberg-Feller central limit theorem (see for example Billingsley (1968), theorem 7.2) yields the desired result.

Since f_n converges to f uniformly and the maximum diameter δ_n of the cells on the whole Δ goes to zero, one can replace f by f_n at a cost going to zero, say ε_n. Since the α_i are independent and have identical distribution, the variance of the sum is the sum of the variances of the random variables $|\Delta_i^n|\, f_i^n(\alpha)$. Hence,

$$\mathrm{Var}\,(n^{1/2}\,\nu_n(f)) = n \sum_{i\geq 1} \left(|\Delta_i^n|^2\, \mathrm{Var}\,(f_i^n(\alpha)) + |\Delta_i^n|\, O(\varepsilon_n)\right).$$

Writing $|\Delta_i^n|^2 \leq n^{-1}\,|\Delta_i^n|$ in the right hand side of this equation, an upper bound of the variance of $n^{1/2}\,\nu_n(f)$ is the mean over Δ of the function $\mathrm{Var}\,(f_n(\cdot,\alpha)) + O(\varepsilon_n)$. Taking only into account the cells Δ_i^n of volume n^{-1}, a lower bound is the mean of $\mathrm{Var}\,(f_n(\cdot,\alpha)) + O(\varepsilon_n)$ over Δ_n^*. The difference between the upper and lower bounds is $O(\varepsilon_n + v_n^*)$, hence the proof is complete.

5. Proof of theorem 2.3

Our goal in section 5.1 is to prove the asymptotic relation (5.1), stated at the end of this section, for $D_\lambda = |\Delta_\lambda^+| - |\Delta_\lambda^+|$.

5.1. Preliminaries on the partitions

Consider the $(n+1)$ lengthes of the intervals delimited by the points x_i, $i \leq n$. They are exchangeable and the points x_i are uniformly distributed on $[0,1]$, hence the measure of Δ_n^+, which is the sum of $i = [n/2]+1$ of them amongst $i + j = (n+1)$ follows a Beta law of parameters (i,j). This law has density

$$\Gamma(i)^{-1}\Gamma(j)^{-1}\Gamma(i+j)x^{i-1}(1-x)^{j-1}\,dx$$

on $\Delta = [0,1]$. The coefficient $\Gamma(i)^{-1}\Gamma(j)^{-1}\Gamma(i+j)$ behaves like 2^n in the exponential scale, hence one can show that $|\Delta_n^+|$ satisfies a large deviations principle

$$P(|\Delta_n^+| \in A) \sim \exp(-n\, c_{\mathrm{DET}}(A))$$

for the action c_{DET} defined by

$$c_{\mathrm{DET}}(y) = -\tfrac{1}{2}\log(4y(1-y))$$

if $y \in (0,1)$ and $c_{\mathrm{DET}}(y) = +\infty$ otherwise. We translate this result on the process indexed by λ. Since the law of $|\Delta_\lambda^+|$ is the barycenter of the laws of $|\Delta_n^+|$ for $n \geq 0$ with coefficients $e^{-\lambda}\lambda^n/n!$, one gets directly

$$P(|\Delta_\lambda^+| \in A) \sim \exp(-\lambda\, c(A))$$

for the action $c = 1 - \exp(-c_{\mathrm{DET}})$. The well known Varadhan's integral lemma (see Dembo and Zeitouni (1992), section 4.3) yields

$$E(\exp(\lambda\, t\, |\Delta_\lambda^+|)) \sim \exp(\lambda\, d(t))$$

for any real t, where d is the Legendre transform of c. We use this result in the following form: first, instead of $|\Delta_\lambda^+|$, we study

$$D_\lambda = |\Delta_\lambda^+| - |\Delta_\lambda^+| = 2\,|\Delta_\lambda^+| - 1.$$

Define d_2 by $d_2(t) = d(2t) - t$. One gets

$$E(\exp(\lambda\, t\, D_\lambda)) \sim \exp(\lambda\, d_2(t))$$

for any real t. Then, d_2 is the Legendre transform of c_2 defined by

$$c_2(2y-1) = c(y).$$

One checks that c_2 is even, $c_2(0) = 0$ and c_2 is infinite outside of $[-1,+1]$. This implies that the function d_2 is even, 1-Lipschitz and $d_2(0) = 0$. These are the only properties of d_2 that we use below.

(Of course, c_2 and d_2 are explicitly known: $c_2(y) = 1 - \sqrt{1-y^2}$ if $|y| \leq 1$ and $c_2(y) = +\infty$ otherwise, $d_2(t) = \sqrt{1+t^2} - 1$ for all t.)

Finally, for any real t and any $\varepsilon > 0$, there exists a finite constant $e(t,\varepsilon)$ such that, for any $\lambda > 0$, one has

$$\lambda\,(d_2(t) - \varepsilon) - e(t,\varepsilon) \leq \log E(\exp(\lambda\, t\, D_\lambda)) \leq \lambda\,(d_2(t) + \varepsilon) + e(t,\varepsilon). \quad (5.1)$$

We choose the coefficient $e(t,\varepsilon)$ even in t so that equation (5.1) holds for $\pm t$ at the same time.

5.2. (H2) and (H3) hold for FRp

Using Section 3.1, we have to check (H2) and (H3) for $f \in C_K$. Writing $g(x)$ for $\frac{1}{2}(f(x,a^+) - f(x,a^-))$, one gets

$$\mu_\lambda(\exp \lambda f) = \exp(\lambda \nu^{1/2}(f))\,\rho_\lambda(\lambda g)$$

with an error term of the form

$$\rho_\lambda(\lambda g) = E\left[\exp\left(\lambda \int_{\Delta_\lambda^+} g - \lambda \int_{\Delta_\lambda^-} g\right)\right].$$

Fix $N \geq 1$ and introduce the partition $\{\Delta_i^N\}_i$ of Δ defined by

$$\Delta_i^N = [(i-1)/N, i/N[.$$

By the uniform continuity of f, there exists a function g_N which is constant on each cell Δ_i^N and such that the uniform norm of $g - g_N$ goes to zero when N goes to infinity. The replacement of g by g_N in the evaluation of $\rho_\lambda(\lambda g)$ causes an error which is $\exp(O(\lambda \|g - g_N\|_\infty))$, hence we can treat g_N only. Denoting by g_i^N the value of g_N on Δ_i^N, one gets

$$\rho_\lambda(\lambda g_N) = E\left[\exp \sum_{i \geq 1} \lambda g_i^N (|\Delta_\lambda^+ \cap \Delta_i^N| - |\Delta_\lambda^- \cap \Delta_i^N|)\right].$$

The restrictions of the Poisson process to each Δ_i^N are i.i.d. Composed by an affine transformation which sends Δ_i^N to Δ, the restriction to Δ_i^N follows the law of the restriction to Δ of a Poisson process of constant intensity λ/N. Hence,

$$|\Delta_\lambda^+ \cap \Delta_i^N| - |\Delta_\lambda^- \cap \Delta_i^N|$$

follows the law of $\pm D_{\lambda/N}/N$, where the $\pm$ sign is measurable with respect to the original Poisson process restricted to $\Delta \setminus \Delta_i^N$: this sign is $+$ if and only if the number of points at the left of Δ_i^N is even. If this number is odd, Δ_i^N begins with an interval of parameter a^- instead of a^+. Call η_i^N this sign and E' the expectation with respect to the law of a copy D' of D, which is independent of everything else. Conditioning by the value of the η's, one has

$$\rho_\lambda(\lambda g_N) = E\left[\prod_{i\geq 1} E'[\exp(\lambda\, g_i^N\, \eta_i^N\, D'_{\lambda/N}/N)]\right].$$

This formula gives estimates of $\rho_\lambda(\lambda g_N)$ as follows. For the upper bound, equation (5.1) yields

$$\log E'[\exp(\lambda\, g_i^N\, \eta_i^N\, D'_{\lambda/N}/N)] \leq (\lambda/N)(d_2(\eta_i^N\, g_i^N) + \varepsilon) + e(\eta_i^N\, g_i^N, \varepsilon).$$

The functions d_2 and $e(\cdot, \varepsilon)$ are even so that we can skip all the η's. Summing up,

$$\log \rho_\lambda(\lambda g_N) \leq \sum_{i\geq 1}\left(\frac{\lambda}{N}(d_2(g_i^N) + \varepsilon) + e(g_i^N, \varepsilon)\right).$$

The right hand side is λ times the integral of $d_2(g_N) + \varepsilon$ over Δ, plus an error term which is constant when N and ε are fixed. This shows

$$\limsup_{\lambda\to\infty} \lambda^{-1} \log \rho_\lambda(\lambda g) \leq 2\,\|g - g_N\|_\infty + \varepsilon + \int_\Delta d_2(g(x))\, dx.$$

Letting ε go to zero and then N go to infinity, one gets the upper bound. The lower bound is proved in the same way. Finally, (H1) holds with

$$l(f) = \nu^{1/2}(f) + \int_\Delta d_2(g(x))\, dx, \quad g = \tfrac{1}{2}(f(\cdot, a^+) - f(\cdot, a^-)).$$

Since d_2 is 1-Lipschitz and $d_2(0) = 0$, $|d_2(y)| \leq |y|$ for all y so that l is finite valued. Since d_2 is differentiable, l is Gateaux differentiable on C_K. The proof will be complete if we identify the Legendre transform i of l.

5.3. Identification of the action

Step 1. The domain of i.

Assume that $i(\nu)$ is finite for a given bounded measure ν. For any $f \in C_K$ and any real t,

$$i(\nu) \geq t(\nu(f) - \nu^{1/2}(f)) - |t|\int_\Delta |g|.$$

Letting t go to $\pm\infty$ yields the following estimations on $\nu(f)$:

$$\nu^{1/2}(f) - \int_\Delta |g| \leq \nu(f) \leq \nu^{1/2}(f) + \int_\Delta |g|.$$

Hence, $\nu(f)$ is always between the means over Δ of the functions

$$\min(f(\cdot, a^+), f(\cdot, a^-)) \quad \text{and} \quad \max(f(\cdot, a^+), f(\cdot, a^-)).$$

This proves that ν is a probability measure and that ν is absolutely continuous with respect to the measure

$$dx\,(\delta_{a^+}(dy) + \delta_{a^-}(dy)).$$

There exists two measurable functions s^+, $s^-: \Delta \to \mathbb{R}^+$ such that

$$\nu(f) = \int_\Delta (s^+(x)\, f(x, a^+) + s^-(x)\, f(x, a^-))\, dx$$

for any measurable bounded function f. Choosing the function $f(x,y) = \mathbf{1}_B(x)$ for any measurable B included in Δ shows that $s^+ + s^- = 1$ a.s. Finally, $i(\nu)$ is finite only when $\nu = \nu^p$ for a given p.

STEP 2. THE VALUE OF i ON ITS DOMAIN.
Assume now that $\nu = \nu^p$ for a measurable $p : \Delta \to [0,1]$. Then,

$$i(\nu^p) = \sup\{\int_\Delta (2p-1)\, g - d_2(g)\,;\, f \in C_K\}.$$

We can compute this supremum directly on $g \in C_K$. Furthermore, since the function $g \mapsto (2p-1)\, g - d_2(g)$ is 1-Lipschitz in L^1 and C_K is dense in L^1 for the L^1 norm, $i(\nu^p)$ is also the supremum over all the functions $g \in L^1$. By the definition of the Legendre transform d_2^* of d_2,

$$(2p(x)-1)\, y - d_2(y) \le d_2^*(2p(x)-1)$$

for any real y. Since d_2 is the Legendre transform of c_2, one gets

$$d_2^*(2p(x)-1) = c_2(2p(x)-1) = c(p(x))$$

and we proved that $i(\nu^p)$ is at most the mean over Δ of $c(p)$. On the other hand, for $\varepsilon \in]0, \frac{1}{2}[$, consider the following truncated approximation p_ε of p and the measurable function g defined as follows:

$$p_\varepsilon = (1-\varepsilon) \wedge (p \vee \varepsilon), \quad g(u) = \frac{(2u-1)}{\sqrt{1-(2u-1)^2}}.$$

Then, $g(u)$ is negative for $u \le 1/2$ and positive otherwise. Furthermore, $(2u-1)\, g(u) - d_2(g(u)) = c(u)$ and $\varepsilon \le \frac{1}{2} \le 1-\varepsilon$, so that

$$\begin{aligned}
[(2p-1)\, g(p_\varepsilon) - d_2(g(p_\varepsilon))] - c(p_\varepsilon) &= 2(p-p_\varepsilon)\, g(p_\varepsilon) \\
&= 2(p-\varepsilon)\mathbf{1}_{p\le\varepsilon} g(\varepsilon) \\
&\quad + 2(p-(1-\varepsilon))\mathbf{1}_{p\ge 1-\varepsilon} g(1-\varepsilon) \\
&\ge 0.
\end{aligned}$$

Since g is bounded, it is integrable and, using g in the definition of $i(\nu^p)$, one gets

$$i(\nu^p) \ge \int_\Delta (2p-1)\, g(p_\varepsilon) - d_2(g(p_\varepsilon)) \ge \int_\Delta c(p_\varepsilon)$$

Letting ε go to zero proves that $i(\nu^p)$ is as stated in Theorem 2.3.

5.4. LARGE DEVIATIONS FOR a_λ

The contraction principle used in Section 3.3 yields

$$P(a_\lambda \in A) \sim \exp(-\lambda\, i_a(A)).$$

The action i_a is as follows. For any measurable admissible $p : \Delta \to [0,1]$, call f^p the function $f^p(x) = p(x)\, a^+(x) + (1-p(x))\, a^-(x)$. Then,

$$i_a(f^p) = i(\nu^p) = \int_\Delta c(p(x))\, dx$$

and $i_a(f) = +\infty$ otherwise. Hence, $i_a(f)$ is the mean over Δ of a local cost at point x given by $j(a^+(x), a^-(x), f(x))$, where the function j is defined as follows:

$$j(x_+, x_-, y) = c\left(\frac{y - x_-}{x_+ - x_-}\right)$$

if $x_+ \neq x_-$, $j(x,x,y) = +\infty$ if $y \neq x$ and $j(x,x,x) = 0$.

5.5. An extension

As we remarked in Section 2 after Theorem 2.5, it is possible to replace the two values $a^\pm$ by two functions $a^\pm : \Delta \to \mathbb{R}$. Assume that $a^\pm$ are a.e. continuous. The proof of Theorem 2.3 can then be copied, except in Step 2 of Section 5.3, when we evaluate the supremum defining $i(\nu^p)$. For any $f \in C_K$, g must be zero on the set

$$\Delta_e = \{x \in \Delta\,;\, a^+(x) = a^-(x)\},$$

hence it is impossible to replace the supremum over f by a supremum over g. However, notice that the choice of $p(x)$ for $x \in \Delta_e$ is irrelevant to the definition of ν^p. This suggests the following way of fixing the problem.

Call admissible any measurable function $p : \Delta \to [0,1]$ such that $p(x) = \frac{1}{2}$ if $x \in \Delta_e$. Then, $i(\nu^p)$ for an admissible p can be evaluated as the supremum over g. The end of the proof is as before.

This modification is effective if $|\Delta_e| > 0$. Then $i(\nu^p) < 1$ for all admissible p since $c(p(x)) = 0$ if $x \in \Delta_e$. In terms of the action i_{DET} of the model FR, this implies that $i_{\mathrm{DET}}(\nu^p)$ is finite for all admissible p, even when ν^p describes a given deterministic material, that is when $p(x)$ is always 0 or 1 on $\Delta \setminus \Delta_e$. This is in contrast with Theorem 2.5 but here is a heuristic explanation of this fact.

Assume that Δ_e has positive measure, that ν^p describes a given deterministic material and that ν^p can be realized as ν_{n_0} at a given cost. If we add an even number of points, say $n - n_0$, in Δ_e, the Young measure is not changed. Starting now from ν_n, we can choose n_0 points amongst n, use them to construct ν_{n_0} which realizes ν^p, and drop the $n - n_0$ other points in Δ_e. In other words, it is at least $C_n^{n_0}\,|\Delta_e|^{n-n_0}$ as easy to realize ν^p with ν_n than with ν_{n_0}, where $C_n^{n_0}$ is the number of ways of choosing n_0 points amongst n and $|\Delta_e|^{n-n_0}$ is the probability that the $n - n_0$ remaining points fall in Δ_e. This suggests that $\exp(-n\, i_{\mathrm{DET}}(\nu^p))$ cannot decrease more rapidly than $C_n^{n_0}\,|\Delta_e|^{n-n_0}$, *i.e.* more rapidly in the exponential scale than $|\Delta_e|^n$ (since $C_n^{n_0}$, for a given n_0, increases as a polynomial in n). As a matter of fact,

$$i_{\mathrm{DET}}(\nu^p) = -\log \int_\Delta \mathbf{1}_{\Delta_e} = -\log|\Delta_e|,$$

so that $|\Delta_e|^n$ is the actual rate of decrease.

6. Proof of theorem 2.4

We begin as usual with results about $D_\lambda = |\Delta_\lambda^+| - |\Delta_\lambda^-|$ and apply them to ν_λ.

6.1. Preliminaries on the partition

Lemma 6.1. *For each $p > 0$, there exists a finite constant b_p such that, for all $n \geq 1$ and $\lambda > 0$,*

$$E(|D_n|^p) \leq b_p\, n^{-p/2}, \quad E(|D_\lambda|^p) \leq b_p\, \lambda^{-p/2}.$$

Furthermore, $n^{1/2}\, D_n$ and $\lambda^{1/2}\, D_\lambda$ converge in law to a centered reduced Gaussian law random variable $\mathcal{N}$.

Proof. We assume throughout that p is an even integer since the result for p implies the result for any $p' \leq p$. From the explicit distribution of Δ_n^+ given in Section 5.1, one can show that the density of the law of D_n on $[-1,+1]$ is bounded, for $n \geq 1$ and up to an absolute constant, by

$$n^{1/2}\,(1-y^2)^{(n/2)-1}.$$

This implies, through classical computations, the existence of an absolute constant κ such that

$$E(|D_n|^p) \leq \kappa\, n^{n/2}\, p^{p/2}\,(n+p)^{-(n+p)/2},$$

for all $n \geq 1$. Now, the coefficient on the right can be bounded:

$$n^{n/2}\,(n+p)^{-(n+p)/2} \leq (n+p)^{-p/2} \leq n!/(n+p/2)!,$$

so that this result applies to D_λ:

$$E(|D_\lambda|^p) \leq \kappa\, p^{p/2}\, e^{-p/2}\, \lambda^{-p/2}.$$

The random variable $n^{1/2}\, D_n$ converges in law to $\mathcal{N}$ by its explicit distribution and by the evaluation of the Gamma function through Stirling's formula.

To deduce the same result for $\lambda^{1/2}\, D_\lambda$, first choose $\varepsilon > 0$. The probability that the number of points of the Poisson process with intensity λ is not in $((1-\varepsilon)\lambda, (1+\varepsilon)\lambda)$ goes to zero exponentially when λ goes to infinity, call this probability $p(\lambda, \varepsilon)$. Hence, for any Borel subset A of $\mathbb{R}$, the difference between $P(\lambda^{1/2}\, D_\lambda \in A)$ and $P(\mathcal{N} \in A)$ is at most $2\, p(\lambda, \varepsilon)$ plus the maximum of the difference between $P(\lambda^{1/2}\, D_n \in A)$ and $P(\mathcal{N} \in A)$ for $n \in ((1-\varepsilon)\lambda, (1+\varepsilon)\lambda)$. Now, $\lambda^{1/2}\, D_n$ is in A if and only if $n^{1/2}\, D_n$ is in $(n/\lambda)^{1/2}\, A$, and it is enough to show that

$$\limsup_{n\to\infty} \sup_{|t|\leq\varepsilon} |P(n^{1/2}\, D_n \in (1+t)\, A) - P(\mathcal{N} \in A)|$$

goes to zero when ε goes to zero. Assume for example that A is an interval (a,b) with $b > a > 0$. Then, the supremum over t can be restricted to the biggest and the smallest sets encountered. Hence, the error is at most the largest of the two numbers

$$\begin{aligned} & P(\mathcal{N} \in ((1-\varepsilon)a, (1+\varepsilon)b)) - P(\mathcal{N} \in (a,b)) \\ \text{and} \qquad & P(\mathcal{N} \in (a,b)) - P(\mathcal{N} \in ((1+\varepsilon)a, (1-\varepsilon)b)), \end{aligned}$$

which both go to zero when ε goes to zero. The other cases are similar. □

6.2. L^p estimates

Assume that $|f|$ is bounded by 1. As usual, we fix N, we replace g by a function g_N which is constant on each interval Δ_i^N and we call g_i^N its value on Δ_i^N. The infinite norm of $g - g_N$ is at most $N^{-\beta}$, hence the replacement of g by g_N in the evaluation of $|\nu_\lambda(f) - \nu^{1/2}(f)|^p$ costs at most $p\,N^{-\beta}$. Now, we work with g_N. We want to compute the p-th power of

$$\tau_\lambda(g_N) = \left|\sum_i g_i^N (|\Delta_\lambda^+ \cap \Delta_i^N| - |\Delta_\lambda^- \cap \Delta_i^N|)\right|.$$

Assume that p is an integer and develop $\tau_\lambda(g_N)^p$ along the multinomial coefficients. Each coefficient $|g_i^N|$ is bounded by 1. By independence and scaling properties of the contributions of each Δ_i^N, one gets

$$E(\tau_\lambda(g_N)^p) \leq \sum_r C_p^r \prod_i E(|D_{\lambda/N}/N|^{r_i}),$$

where r is any integer-valued multi-index of size N and sum p and C_p^r is the multinomial coefficient of r. Denote by c_p the maximum of the constants b_q in Lemma 6.1 for $q \leq p$. Notice that at most p terms in the multi-index r are not zero. Lemma 6.1 yields

$$E(\tau_\lambda(g_N)^p) \leq c_p^p N^{-p} \sum_r C_p^r \prod_i (\lambda/N)^{-r_i/2} = c_p^p (\lambda/N)^{-p/2}.$$

Finally, the expectation of $|\nu_\lambda(f) - \nu^{1/2}(f)|^p$ is at most $N^{-\beta} + (\lambda/N)^{-p/2}$. If N is equivalent to $\lambda^{p/(p+2\beta)}$, the error term is equivalent to $\lambda^{-p\beta/(p+2\beta)}$.

6.3. Central limit theorem

Call $(\Delta_i)_i$ a finite partition of Δ, adapted to f. This means that f is constant on each set Δ_i. Call g_i the value of g on Δ_i. The independence and scaling properties of the restrictions of the Poisson process to the sets Δ_i give

$$\lambda^{1/2}(\nu(f) - \nu^{1/2}(f)) = \sum_{i \geq 1} \lambda^{1/2} g_i \eta_i L_i D_{\lambda L_i}^{(i)},$$

where L_i is the Lebesgue measure of Δ_i, $\eta_i = \pm 1$ and the $D^{(i)}$ are independent Poisson processes (see the decomposition of Section 5.2). Each $(\lambda L_i)^{1/2}\, D_{\lambda L_i}^{(i)}$ converges to the reduced Gaussian law. Furthermore, if a Poisson process is conditioned to have an even number of points (or an odd number of points) and if D is constructed with this conditioned Poisson process, then $\lambda^{1/2}\, D_\lambda$ converges also to the centered Gaussian law. Hence, conditioned by the η's, $(\lambda L_i)^{1/2}\, D_{\lambda L_i}^{(i)}$ converges to a Gaussian reduced law $\mathcal{N}_i$ which is independent of the η's. Finally, conditioned by the value of $(\eta_i)_i$, the random variable $\lambda^{1/2}(\nu(f) - \nu^{1/2}(f))$ converges to

$$\sum_{i \geq 1} g_i L_i^{1/2} \eta_i \mathcal{N}_i$$

where the $\mathcal{N}_i$ are independent of the η's, i.i.d. and reduced Gaussian. Whatever the value of $(\eta_i)_i$ is, this sum is Gaussian, centered of variance

$$\sum_{i \geq 1} (g_i L_i^{1/2} \eta_i)^2 = \int_\Delta g^2(x)\, dx,$$

and this proves Theorem 2.4.

6.4. Model FR

In order to prove the L^p estimate, we develop as before $\tau_n(g_N)^p$ as a multinomial when p is an integer. Call n_i the number of points in Δ_i^N and condition by $(n_i)_i$. Things become independent, one can use Lemma 6.1 and it gives an upper bound:

$$E(\tau_n(g_N)^p) \leq c\, N^{-p} \sum_r C_p^r\, E\left(\prod_{i=1}^{N} (n_i^{-r_i/2} \wedge 1)\right).$$

We fix r and we want to bound the expectations written in the right hand side. The mean of $(n_i)_i$ is the point $(nN^{-1})_{i\leq N}$ in $\mathbb{R}^N$. We split the expectations in two parts. On the ball around $(nN^{-1})_{i\leq N}$ of radius $R = \frac{1}{2}nN^{-1}$, the random variable is at most $(2Nn^{-1})^{p/2}$. On the complement of the ball, the random variable is at most 1 but the probability that $(n_i)_i$ is in the complement of the ball is, by Bienaymé-Tchebychev inequality, at most

$$\mathrm{Var}\,((n_i)_i)\, R^{-2} \leq 4N^2\, n^{-1}.$$

Finally, each expectation is at most $O(N^2\, n^{-1} + N^{p/2}\, n^{-p/2})$ and the sum of the C_p^r is N^p. The error due to the replacement of g by g_N is at most $O(N^{-\beta})$, hence one has to get

$$N^{-\beta} + N^2\, n^{-1} + N^{p/2}\, n^{-p/2}$$

as small as possible, that is to choose $N \sim n^{1/(2+\beta)}$.

Notice that the rate for the model FRp is always better (for $p \geq 1$) than the rate for the model FR.

Appendix: Large deviations for FR

Although the proof of the large deviations principle for the model FR follows the same general line, it is much more intricate than for the model FRp. It would be interesting to deduce directly the result from its analog for model FRp.

An auxiliary result

We first need a minor modification of a well known result. Varadhan's integral lemma (see for example Dembo and Zeitouni (1992), section 4.3) states that, if μ_n satisfies a large deviations principle of action i_0 and if f is a bounded continuous function, then $n^{-1} \log \mu_n(\exp nf)$ converges to the supremum of $(f - i_0)$. Now, assume that i_0 is continuous on its domain and consider a nonnegative function f', not necessarily bounded nor continuous, such that the interior of the support of f' is empty. Then,

Lemma A.1. *(Modified Varadhan's lemma)* $n^{-1} \log \mu_n(\exp n(f - f'))$ *converges to the supremum of* $(f - i_0)$.

Proof. Since f' is nonnegative, the upper bound is direct. For the lower bound, denote by S the support of f' and choose $x \notin S$ in the domain of i_0. Since S is closed, there exists an open ball B around x of radius $\varepsilon > 0$, which is disjoint from S. Taking only into account the integral over B, one gets

$$\mu_n(\exp n(f - f')) \geq \mu_n(B)\exp(n \inf_B f).$$

Take the logarithm, divide by n and take the liminf as n goes to infinity. One gets

$$\liminf_{n\to\infty} n^{-1} \log \mu_n(\exp n(f - f')) \geq -i_0(B) + \inf_B f.$$

Take the limit as ε goes to zero. Since f is continuous and i_0 is lower semicontinuous, the lower bound converges to $f(x) - i_0(x)$. It remains to show that the supremum of $f - i_0$ on the intersection of the complement of S with the domain of i_0 is in fact the supremum of $f - i_0$. This stems from the fact that $f - i_0$ is continuous, that $f(x) - i_0(x) = -\infty$ if x is not in the domain of i_0 and that the complement of S is a dense subset. □

The law of D_n

Next, we prove precise estimates for the law of D_n. We know that $E(\exp(ntD_n))$ is exponentially equivalent to $\exp(n\,j(t))$ where j is the Legendre transform of the function

$$y \mapsto c_{\mathrm{DET}}\left(\tfrac{1}{2}(1+y)\right) = -\tfrac{1}{2}\log(1-y^2)$$

if $|y| < 1$ and $+\infty$ otherwise. Since j is the Legendre transform of an even function which is null at zero and infinite outside of $|y| < 1$, j is even, convex, null at zero, increasing on $t \geq 0$ and 1-Lipschitz. Hence, $j(t) \leq |t|$. An explicit form of $j(t)$ is as follows:

$$j(t) = \tfrac{1}{2}[s - 1 - \log(\tfrac{1}{2}(s+1))], \quad s = \sqrt{1+4t^2}.$$

Hence, $j(t)$ is equivalent to $|t|$ when $|t|$ goes to infinity and we extend the function $t\,j(c/t)$ by the value $|c|$ when $t = 0$. Explicit computations show that the function j_c defined by

$$j_c(t) = t\,j(c/t) - t\,\log t$$

is concave on $t \geq 0$. The estimates of $E(e^{n\,t\,D_n})$ are summarized in the next lemma.

Lemma A.2. *There exist even functions $e_\pm(n,\cdot) \geq 0$ such that, for all t and $n \geq 1$,*

$$n\,j(t) - e_-(n,t) \leq \log E(e^{n\,t\,D_n}) \leq n\,j(t) + e_+(n,t).$$

There exist $c > 0$ and $t_* > 0$ such that the following bounds hold.
– For all $n \geq 0$ and t, one has: $e_+(n,t) \leq c$.
– For all $n \geq 1$ and t, one has: $e_-(n,t) \leq 2|t| + 4\log(n) + c$.
– For all $n \geq 1$ and $|t| \geq t_*$, one has:

$$e_-(n,t) \leq 3n\,t^{-2} + 3\log|t| + \log(n) + c.$$

– For any t, one has: $-|t| \leq \log E(\exp(t\,D_0)) \leq |t|$.

Proof. The following is only a sketch of the proof. Use the law of D_n to write the exact value of the desired expectation. Apply Stirling's formula to get upper and lower bounds of $\Gamma(i)^{-1}\,\Gamma(j)^{-1}\,\Gamma(i+j)$, for the right values of i and j, by a power of n. Remains the integral of an exponential. Then copy the classical proof of Laplace's formula.

More precisely, get an upper bound by replacing the exponent by its maximum. To get the lower bound which is true for all t, keep only the integral on an interval of length n^{-1} around the point x_t where the exponent is maximal, and replace the exponent by its minimum on this interval. To get the asymptotic lower bound, keep only the integral on an interval of length t^{-2} around the point x_t and proceed in the same way. The cases n even and n odd are to be treated separately (the case n odd is simpler because the law of D_n is then symmetric).
To get the bound involving D_0, notice that $D_0 = 1$. □

The condition (H2)

We are now ready to check the conditions (H2) and (H3) for model FR. The beginning of the proof is similar to the case FRp. For any $f \in C_K$,

$$E(\exp \nu_n(nf)) = \exp(O(n\,\|g - g_N\|_\infty))\,\exp(\nu^{1/2}(nf))\,\rho_n(n\,g_N)$$

where g_N is a function which is constant on every interval Δ_i^N and close to g in uniform norm, and

$$\rho_n(n\,g_N) = E\left[\exp\sum_i n\,g_i^N\,(|\Delta_i^N \cap \Delta_n^+| - |\Delta_i^N \cap \Delta_n^-|)\right].$$

Call n_i the number of points of the uniform sample which belong to Δ_i^N. Conditioned by n_i, these points are a uniform sample of size n_i taken from Δ_i^N and independent of the rest of the process. Hence, conditioning by $(n_i)_i$, there exist signs $\eta_i^N = \pm 1$ and random variables D' independent of the η's such that

$$\rho_n(n\,g_N) = E\left[\prod_i E'(\exp(n\,g_i^N\,\eta_i^N\,D'_{n_i}/N)\right]$$

where E' denotes the expectation with respect to D'. Hence, each expectation E' is really a random variable which is η_i^N measurable. We apply Lemma A.2 to each expectation E'. Even when $n_i = 0$, the principal term is

$$\exp[n_i\,j(n\,g_i^N\,\eta_i^N/(N\,n_i))].$$

Since the function j is even, the principal term for $\rho_n(ng_N)$ is

$$\widehat{\rho_n}(ng_N) = E\left(\exp\left(\sum_i n_i\,j(n\,g_i^N\,\eta_i^N/(N\,n_i))\right)\right).$$

From now on, we compute upper and lower bounds of $\rho_n(ng_N)/\widehat{\rho_n}(ng_N)$, estimating the contribution of each expectation E' to the overall error. For the upper bound, the error due to each expectation E' is at most $\exp c$, even when $n_i = 0$. Hence, $\rho_n(ng_N)/\widehat{\rho_n}(ng_N)$ is at most $\exp(N\,c) = \exp(o(n))$. (Remember that N is a fixed integer and that n goes to infinity.)

The lower bound is more difficult to handle. We want an upper bound of $\widehat{\rho_n}(ng_N)/\rho_n(ng_N)$. First, since the bounds are even, we can skip the η's. Furthermore, since the bounds are increasing in $|t|$, we can replace g_i^N by 1. Hence, we apply Lemma A.2 for n_i and $t_i = n/(N\,n_i)$.

– If $n_i \geq n^{1/2}$, then $t_i \leq n^{1/2}/N$ and the bound of e_- which is valid for all t gives an error
$$c\, n_i^4 \exp(2n^{1/2}/N).$$
There are N expectations E', hence the total error due to such terms is at most
$$c^N\, n^{4N} \exp(2n^{1/2}) = \exp(o(n)).$$
– If $g_i^N = 0$, no error term.
– If $1 \leq n_i \leq n^{1/2}$ and $g_i^N \neq 0$, we want to use the asymptotic lower bound of Lemma A.2. Hence, we set g_0^N equal to the least absolute value of the non zero g_i^N and we assume that $n^{1/2} g_0^N \geq t_* N$. Then $t_i = |g_i^N|\, n/(N\, n_i)$ is at least t_*, hence the error is at most
$$cn_i\, |t_i|^3 \exp(3n_i/t_i^2) \leq cn\, (n/N)^3 \exp(3(N/g_0^N)^2)$$
because g^N is bounded. There are at most N terms, hence the total error due to the terms of this kind is $\exp(o(n))$.
– Last, if $n_i = 0$, the error is at most $\exp(2|t_i|)$ for $t_i = n\, g_i^N/N$, that is at most $\exp(2n/N)$.

All the errors disappear, except the last one. The liminf of $n^{-1} \log \rho_n(ng_N)$ is at least the liminf of
$$n^{-1} \log E \left(\exp \left\{ \sum_i n_i\, j(n\, g_j^N/(N\, n_i)) - 2n\, \mathbf{1}_{\{n_i=0\}}/N \right\} \right).$$
The random vector $(n_i)_i$ is the sum of n i.i.d. random vectors distributed like ξ: the random vector ξ is uniformly distributed on the vectors of the canonical orthonormal basis of $\mathbb{R}^N$. Since ξ is bounded, $(n_i/n)_i$ satisfies a large deviations principle of continuous action i_0. In fact, i_0 is the Cramér transform of the law of ξ:
$$i_0(x) = \sup\{t.x - \log E(\exp(t.\xi))\,;\, t \in \mathbb{R}^N\}.$$
The explicit value of i_0 is
$$i_0(x) = \log(N) + \sum_i x_i \log(x_i)$$
if $x = (x_i)_i$, $x_i \geq 0$ and $\sum_i x_i = 1$, and $i_0(x) = +\infty$ otherwise. We set
$$f_N(x) = \sum_i x_i\, j(g_i^N/(Nx_i))$$
$$f'_N(x) = 2N^{-1} \sum_i \mathbf{1}(x_i = 0)$$
for $x = (x_i)_i$ and $x_i \geq 0$. By lemma A.1, f'_N can be omitted without changing the limit and the limit of this quantity, hence of $n^{-1} \log \rho_n(n\, g_N)$ as well, is the supremum of $f_N - i_0$.

Define a function h_N on Δ which is constant on each Δ_i^N by setting $h_i^N = Nx_i$. Then, the mean of h_N is 1, h_N is nonnegative, $f_N(x)$ is the mean of $h_N\, j(g_N/h_N)$ over Δ and $i_0(x)$ is the mean of $h_N \log h_N$ over Δ. Finally, the 1-Lipschitz regularity of j is enough to show that (H2) holds with
$$l_{\mathrm{DET}}(f) = \nu^{1/2}(f) + \sup\{ \int_\Delta h\, j(g/h) - h \log h\,;\, h \in H\}$$

where H is, for example, the space of integrable, nonnegative measurable functions of mean 1. For $h \in H$, the mean of $h \log h$ is nonnegative and $h\,j(g/h) \le |g|$, hence l_{DET} is finite valued.

The differentiability of l_{DET}

The proof of (H3) is complete if we show the Gateaux differentiability of l_{DET}, or of the function $f \mapsto l_{\mathrm{DET}}(f) - \nu^{1/2}(f)$. For any g, set

$$m(g) = \sup\{m(g,h), h \in H\}, \qquad m(g,h) = \int_\Delta h\,j(g/h) - h \log h.$$

The function $h \mapsto h\,j(g/h) - h\log h$ is concave. If the Gateaux differential of $m(g,\cdot)$ at $h \in H$ is zero in all the directions, then $m(g,h) = m(g)$. Choose the unique $c(g) = c \ge 1$ such that

$$\int_\Delta h_c(x) = 1, \qquad h_c(x) = \frac{c^2}{\sqrt{c^2 + g^2(x)}}.$$

A small amount of algebra shows that the differential of $m(g,\cdot)$ is null at h_c. Another amount of algebra gives

$$\begin{aligned} m(g) &= m(g, h_c) \\ &= -1 - \log c(g) - \int_\Delta \sqrt{c(g)^2 + g^2(x)}\,dx, \end{aligned}$$

and the differentiability of $g \mapsto c(g)$ implies the differentiability of m.

At this moment, we already know that ν_n satisfies a large deviations principle of action i_{DET}. Hence, the function i_{DET} must satisfy

$$\sum_{n \ge 0} e^{-\lambda}\,\lambda^n \exp(-n\, i_{\mathrm{DET}}(\nu))/n! \sim \exp(-\lambda\, i(\nu))$$

when λ goes to infinity, that is $\exp(-i_{\mathrm{DET}}(\nu)) - 1 = -i(\nu)$. In other words, the identification of the Legendre transform of l_{DET}, although possible to accomplish, may be avoided.

References

Baldi, P. (1988) Large deviations and stochastic homogenization. *Ann. Mat. Pura Appl.* **151,** 161–177.

Billingsley, P. (1968) *Convergence of probability measures.* John Wiley & Sons, New York.

Breiman, L. (1968) *Probability.* Addison-Wesley,.

Dal Maso, G. (1993) *An introduction to* Γ*-convergence.* Birkhäuser,.

Dal Maso, G., Modica, L. (1986) Nonlinear stochastic homogenization. *Ann. Mat. Pura Appl.* **144,** 347–389.

Dembo, A., Zeitouni, O. (1992) *Large deviations techniques and applications.* Jones and Bartlett,.

Facchinetti, G., Russo, L. (1983) Un caso unidimensionale di omogeneizzazione stocastica. *Boll. U.M.I. Serie VI* **98,** 159–170.

Jikov (Zhikov), V., Kozlov, S., Oleinik, O. (1994) *Homogenization of differential operators and integral functionals.* Springer, Berlin.

Kozlov, S. (1980) Averaging of random operators. *Math. USSR Sbornik* **37,** 167–180.

Michel, J., Robert, R. (1994) Large deviations for Young measures and statistical mechanics of infinite dimensional dynamical systems with conservation law. *Comm. Math. Phys.* **159,** 195–215.

PIAU, D. (1998) Large deviations and Young measures for a Poissonian model of bi-phased material. submitted.

YURINSKIĬ, V. (1991) Homogenization error estimates for random elliptic operators. *Lect. in Appl. Maths.* **27**, 285–290.

JULIEN MICHEL: UNITÉ DE MATHÉMATIQUES PURES ET APPLIQUÉES (UMR 128), ENS–LYON, 46 ALLÉE D'ITALIE, 69364 LYON CEDEX 07, FRANCE. E-MAIL: jmichel@umpa.ens-lyon.fr

DIDIER PIAU: LABORATOIRE DE PROBABILITÉS, UNIVERSITÉ CLAUDE BERNARD–LYON I, 43 BOULEVARD DU 11 NOVEMBRE 1918, 69622 VILLEURBANNE CEDEX, FRANCE. E-MAIL: piau@jonas.univ-lyon1.fr