A revised version of this paper has been published in
***The Annals of Probability*, 44(1):444-478, 2016.**

A Curie-Weiss Model of Self-Organized Criticality

Raphaël Cerf
Université Paris Sud and IUF

and

Matthias Gorny
Université Paris Sud

Abstract

We try to design a simple model exhibiting self-organized criticality, which is amenable to a rigorous mathematical analysis. To this end, we modify the generalized Ising Curie-Weiss model by implementing an automatic control of the inverse temperature. For a class of symmetric distributions whose density satisfies some integrability conditions, we prove that the sum S_n of the random variables behaves as in the typical critical generalized Ising Curie-Weiss model. The fluctuations are of order $n^{3/4}$ and the limiting law is $C\exp(-\lambda x^4)\,dx$ where C and λ are suitable positive constants.

AMS 2010 subject classifications: 60F05 60K35.
Keywords: Ising Curie-Weiss, self-organized criticality, Laplace's method.

1 Introduction

In their famous article [4], Per Bak, Chao Tang and Kurt Wiesenfeld showed that certain complex systems are naturally attracted by critical points, without any external intervention. The amplification of small internal fluctuations can lead to a critical state and cause a chain reaction leading to a radical change of the system behavior. These systems exhibit the phenomenon of self-organized criticality (SOC). Although there is no universal SOC theory, it can be well understood with the archetype of SOC : the sandpile model, first introduced in [5]. We consider a pile of sand and the constant drop of new sand grains, which randomly slide down the slope of sand. We observe local avalanches with different and unpredictable sizes which are not proportional to the input. Such phenomenon can be observed in nature (e.g., forest fires, earthquakes, species evolution).

In general SOC can be observed empirically or simulated on a computer in various models. However the mathematical analysis of these models turns out to be extremely difficult, even for the sandpile model whose definition is yet simple. Self-organized criticality has been reviewed in recent works [2,3,11,19,23]. Other challenging models are the models for forest fires [20], which are built with the help of percolation process. Some simple models of evolutions also lead to critical behaviours [9].

Our goal here is to design a model exhibiting self-organized criticality, which is as simple as possible, and which is amenable to a rigorous mathematical analysis. The simplest models exhibiting SOC are obtained by forcing standard critical transitions into a self-organized state (see section 15.4.2 of [22]). The idea is to start with a model presenting a phase transition and to create a feedback from the configuration to the control parameters in order to converge towards a critical point. The most widely studied model in statistical mechanics, which exhibits a phase transition and presents critical states, is the Ising model. Its mean field version is called the Ising Curie-Weiss model (see sections IV.4 and V.9 of [13]). It has been extended to real-valued spins by Richard S. Ellis and Charles M. Newman [14], in the so called generalized Ising Curie-Weiss model. This model is our starting point and we will modify it in order to build a system of interacting random variables, which exhibits a phenomenon of SOC.

Let us first recall the definition and some results on the generalized Ising Curie-Weiss model. Let ρ be a symmetric probability measure on $\mathbb{R}$ with positive variance σ^2 and such that

$$\forall t \geq 0 \qquad \int_{\mathbb{R}} \exp(tx^2)\, d\rho(x) < \infty.$$

The generalized Ising Curie-Weiss model associated to ρ and the inverse temperature $\beta > 0$ is defined through an infinite triangular array of real-valued random variables $(X_n^k)_{1\leq k\leq n}$ such that, for all $n \geq 1$, $(X_n^1, \dots, X_n^n)$ has the distribution

$$d\mu_{n,\rho,\beta}(x_1, \dots, x_n) = \frac{1}{Z_n(\beta)} \exp\left(\frac{\beta}{2}\frac{(x_1 + \cdots + x_n)^2}{n}\right) \prod_{i=1}^{n} d\rho(x_i),$$

where $Z_n(\beta)$ is a normalization. For any $n \geq 1$, we set $S_n = X_n^1 + \cdots + X_n^n$. When $\rho = (\delta_{-1} + \delta_1)/2$, we recover the classical Ising Curie-Weiss model.

We denote by L the Log-Laplace of ρ (see appendix A). Richard S. Ellis and Theodor Eisele have shown in [12] that, if $L^{(3)}(t) \leq 0$ for any $t \geq 0$, then there exists a map m which is null on $]0, 1/\sigma^2]$, real analytic and positive on $]1/\sigma^2, +\infty[$, and such that

$$\frac{S_n}{n} \xrightarrow[n\to\infty]{\mathcal{L}} \begin{cases} \delta_0 & \text{if} \quad \beta \leq 1/\sigma^2 \\ \frac{1}{2}(\delta_{-m(\beta)} + \delta_{m(\beta)}) & \text{if} \quad \beta > 1/\sigma^2. \end{cases}$$

The point $1/\sigma^2$ is a critical value and the function m cannot be extended analytically around $1/\sigma^2$. The main result of [14] states that, if $\beta < 1/\sigma^2$, then, under $\mu_{n,\rho,\beta}$,

$$\frac{S_n}{\sqrt{n}} \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}\left(0, \frac{\sigma^2}{1-\beta\sigma^2}\right).$$

If $\beta = 1/\sigma^2$, then there exists $k \in \mathbb{N}\backslash\{0,1\}$ and $\lambda > 0$ such that, under $\mu_{n,\rho,\beta}$,

$$\frac{S_n}{n^{1-1/2k}} \xrightarrow[n\to\infty]{\mathcal{L}} C_{k,\lambda} \exp\left(-\lambda \frac{s^{2k}}{(2k)!}\right) ds,$$

where $C_{k,\lambda}$ is a normalization. This is a consequence of theorem 2.1 of [14] and some properties of m explained in [12] implying that $s \longmapsto L(s\sqrt{\beta}) - s^2/2$ has a unique maximum at 0 whenever $\beta \leq 1/\sigma^2$ (see section V.2 of [16] for the details).

We will transform the previous probability distribution in order to obtain a model which presents a phenomenon of self-organized criticality, i.e., a model which evolves towards the critical state $\beta = 1/\sigma^2$ of the previous model. More precisely, the critical generalized Ising Curie-Weiss model is the model where $(X_n^1, \ldots, X_n^n)$ has the distribution

$$\frac{1}{Z_n} \exp\left(\frac{(x_1 + \cdots + x_n)^2}{2n\sigma^2}\right) \prod_{i=1}^n d\rho(x_i).$$

We wish to build a model which converges to a critical state for every distribution ρ and which does not rely on any specific a priori information on ρ. We search an automatic control of the inverse temperature β, which would be a function of the random variables in the model, so that, when n goes to $+\infty$, β converges towards the critical value of the model. We start with the following observation: if $(Y_n)_{n\geq 1}$ is a sequence of independent random variables with identical distribution ρ, then, by the law of large numbers,

$$\frac{Y_1^2 + \cdots + Y_n^2}{n} \xrightarrow[n\to\infty]{} \sigma^2 \quad \text{a.s.}$$

This convergence provides us with an estimator of $1/\sigma^2$. If we believe that a similar convergence holds in the generalized Ising Curie-Weiss model, then we are tempted to « replace β by $n\,(x_1^2 + \cdots + x_n^2)^{-1}$ » in the distribution

$$\frac{1}{Z_n} \exp\left(\frac{\beta}{2} \frac{(x_1 + \cdots + x_n)^2}{n}\right) \prod_{i=1}^n d\rho(x_i).$$

Hence the model we consider in this paper is given by the distribution

$$\frac{1}{Z_n} \exp\left(\frac{1}{2} \frac{(x_1 + \cdots + x_n)^2}{x_1^2 + \cdots + x_n^2}\right) \prod_{i=1}^n d\rho(x_i).$$

The previous considerations suggest that this model should evolve spontaneously towards a critical state. We will prove rigorously that our model indeed exhibits a phenomenon of self-organized criticality. However our model is a toy model which is certainly much less complex than other famous fundamental models of SOC like the sandpile model.

Our main result (theorem 2) states that, if ρ has an even density satisfying some integrability condition, then, asymptotically, the sum S_n of the random variables behaves as in the typical critical generalized Ising Curie-Weiss model: if μ_4 denotes the fourth moment of ρ, then

$$\frac{\mu_4^{1/4} S_n}{\sigma^2 n^{3/4}} \xrightarrow[n\to\infty]{\mathcal{L}} \left(\frac{4}{3}\right)^{1/4} \Gamma\left(\frac{1}{4}\right)^{-1} \exp\left(-\frac{s^4}{12}\right)\, ds.$$

This fluctuation result shows that our model is a self-organized model exhibiting critical behaviour. Indeed it has the same behaviour than the critical generalized Ising Curie-Weiss model and, by construction, it does not depend on any external parameter. In this sense, we can conclude that this is a Curie-Weiss model of self-organized criticality.

Our result presents an unexpected universal feature. For any distribution ρ, which has an even density satisfying some integrability hypothesis, the fluctuations of S_n are of order $n^{3/4}$. This is in contrast to the situation in the critical generalized Ising Curie-Weiss model: at the critical point, the fluctuations are of order $n^{1-1/2k}$, where k depends on the distribution ρ. We stress also that our integrability conditions on ρ are weaker than those of [14]. For instance, our result holds for any centered Gaussian measure on $\mathbb{R}$. The Gaussian case of our model can be handled with the help of an explicit computation [17].

The main new technical ingredient of the proof is the following inequality. Let Z be a random variable with distribution ρ, and let I denote the Cramér transform of (Z, Z^2), given by

$$\forall (x,y)\in\mathbb{R}^2 \qquad I(x,y) = \sup_{(u,v)\in\mathbb{R}^2} \left\{ xu + yv - \ln \int_{\mathbb{R}} e^{uz+vz^2}\, d\rho(z) \right\}.$$

If ρ is symmetric and there exists $v>0$ such that $E(\exp(vZ^2)) < +\infty$, then

$$\forall (x,y)\in\mathbb{R}^2 \qquad I(x,y) \ \geq\ \frac{x^2}{2y},$$

and the equality holds only at $(0,\sigma^2)$. We explain in the heuristics at the end of section 3 why this inequality is crucial to the proof of our main results.

In section 2 we define properly our model. We state our main results and the strategy for proving them in section 3. Next we split the proofs in the remaining sections (4-7). In appendix, we recall some generalities on the Cramér transform and large deviations.

Acknowledgements. We thank two anonymous Referees for their comments which helped to improve the presentation of the paper.

2 The model

Let ρ be a probability measure on $\mathbb{R}$, which is not the Dirac mass at 0. We consider an infinite triangular array of real-valued random variables $(X_n^k)_{1\leq k\leq n}$ such that for all $n \geq 1$, $(X_n^1,\dots,X_n^n)$ has the distribution $\widetilde{\mu}_{n,\rho}$, where

$$d\widetilde{\mu}_{n,\rho}(x_1,\dots,x_n) = \frac{1}{Z_n}\exp\left(\frac{1}{2}\frac{(x_1+\cdots+x_n)^2}{x_1^2+\cdots+x_n^2}\right)\mathbb{1}_{\{x_1^2+\cdots+x_n^2>0\}}\prod_{i=1}^n d\rho(x_i),$$

with

$$Z_n = \int_{\mathbb{R}^n}\exp\left(\frac{1}{2}\frac{(x_1+\cdots+x_n)^2}{x_1^2+\cdots+x_n^2}\right)\mathbb{1}_{\{x_1^2+\cdots+x_n^2>0\}}\prod_{i=1}^n d\rho(x_i).$$

We define $S_n = X_n^1+\cdots+X_n^n$ and $T_n = (X_n^1)^2+\cdots+(X_n^n)^2$.

The indicator function in the density of the distribution $\widetilde{\mu}_{n,\rho}$ helps to avoid any problem of definition if $\rho(\{0\})$ is positive, since, if $\rho(\{0\}) > 0$, the event $\{x_1^2+\cdots+x_n^2 = 0\}$ may occur with positive probability. We notice that, unlike the generalized Ising Curie-Weiss model, our model is defined for any probability measure. Indeed $x \longmapsto x^2$ is a convex function, therefore

$$\forall (x_1,\dots,x_n)\in\mathbb{R}^n \qquad \left(\sum_{i=1}^n x_i\right)^2 = n^2\left(\sum_{i=1}^n \frac{x_i}{n}\right)^2 \leq n\sum_{i=1}^n x_i^2.$$

Thus for any $n\geq 1$, $1 \leq Z_n \leq e^{n/2} < +\infty$.

If we choose $\rho = (\delta_{-1}+\delta_1)/2$, we obtain the classical Ising Curie-Weiss model at the critical value.

3 Convergence theorems

We state here our main results.

By the classical law of large numbers, if ρ is centered and has variance σ^2, then, under $\rho^{\otimes n}$, $(S_n/n, T_n/n)$ converges in probability towards $(0,\sigma^2)$. The next theorem shows that, under the law $\widetilde{\mu}_{n,\rho}$, given certain conditions, $(S_n/n, T_n/n)$ also converges in probability to $(0,\sigma^2)$.

Theorem 1. *Let ρ be a symmetric probability measure on $\mathbb{R}$ with positive variance σ^2 and such that*

$$\exists v_0 > 0 \qquad \int_{\mathbb{R}} e^{v_0 z^2}\, d\rho(z) < +\infty.$$

We suppose that one of the following conditions holds:

(a) *ρ has a density.*

(b) *ρ is the sum of a finite number of Dirac masses.*

(c) *There exists $c > 0$ such that $\rho(]0,c[) = 0$.*

(d) *$\rho(\{0\}) < 1/\sqrt{e}$.*

Then, under $\widetilde{\mu}_{n,\rho}$, $(S_n/n, T_n/n)$ converges in probability towards $(0,\sigma^2)$.

By the classical central limit theorem, under $\rho^{\otimes n}$, $S_n/\sqrt{n}$ converges in distribution to a normal distribution with mean zero and variance σ^2. The following theorem, shows that, given certain conditions, under $\widetilde{\mu}_{n,\rho}$, $S_n/n^{3/4}$ converges towards a specific distribution.

Theorem 2. *Let ρ be a probability measure on $\mathbb{R}$ with a density f satisfying:*

(a) *f is even.*

(b) *There exists $v_0 > 0$ such that*

$$\int_{\mathbb{R}} e^{v_0 z^2} f(z)\, dz < +\infty.$$

(c) *There exists $p \in]1,2]$ such that*

$$\int_{\mathbb{R}^2} f^p(x+y) f^p(y) |x|^{1-p}\, dx\, dy < +\infty.$$

Let σ^2 be the variance of ρ and let μ_4 be the fourth moment of ρ. We have

$$\frac{\mu_4^{1/4} S_n}{\sigma^2 n^{3/4}} \xrightarrow[n\to\infty]{\mathcal{L}} \left(\frac{4}{3}\right)^{1/4} \Gamma\left(\frac{1}{4}\right)^{-1} \exp\left(-\frac{s^4}{12}\right)\, ds.$$

The convergence can equivalently be rewritten as

$$\frac{S_n}{n^{3/4}} \xrightarrow[n\to\infty]{\mathcal{L}} \left(\frac{4\mu_4}{3\sigma^8}\right)^{1/4} \Gamma\left(\frac{1}{4}\right)^{-1} \exp\left(-\frac{\mu_4}{12\sigma^8} s^4\right)\, ds.$$

We prove this convergence in section 7.

The following corollary is a version of theorem 2 with an hypothesis which is weaker but easier to check.

Corollary 3. *Let ρ be a probability measure on $\mathbb{R}$ with an even and bounded density f such that*

$$\exists\, v_0 > 0 \qquad \int_{\mathbb{R}} e^{v_0 z^2}\, d\rho(z) < +\infty.$$

Let σ^2 be the variance of ρ and let μ_4 be the fourth moment of ρ. Then

$$\frac{\mu_4^{1/4} S_n}{\sigma^2 n^{3/4}} \xrightarrow[n\to\infty]{\mathcal{L}} \left(\frac{4}{3}\right)^{1/4} \Gamma\left(\frac{1}{4}\right)^{-1} \exp\left(-\frac{s^4}{12}\right)\, ds.$$

Proof. We check that the hypothesis of the corollary imply the condition (c) of theorem 2. We have

$$\begin{aligned}
&\int_{\mathbb{R}^2} f^{3/2}(x+y) f^{3/2}(y) |x|^{-1/2}\, dx\, dy \\
&\quad = \int_{[-1,1]\times\mathbb{R}} \frac{f^{3/2}(x+y) f^{3/2}(y)}{|x|^{1/2}}\, dx\, dy + \int_{[-1,1]^c\times\mathbb{R}} \frac{f^{3/2}(x+y) f^{3/2}(y)}{|x|^{1/2}}\, dx\, dy \\
&\quad \le \|f\|_\infty^{3/2} \int_{[-1,1]\times\mathbb{R}} \frac{f^{3/2}(y)}{|x|^{1/2}}\, dx\, dy + \int_{[-1,1]^c\times\mathbb{R}} f^{3/2}(x+y) f^{3/2}(y)\, dx\, dy \\
&\quad \le \|f\|_\infty^{3/2} \left(\int_{\mathbb{R}} |f(x)|^{3/2}\, dx\right) \left(\int_{-1}^{1} \frac{dx}{|x|^{1/2}}\right) + \left(\int_{\mathbb{R}} |f(x)|^{3/2}\, dx\right)^2.
\end{aligned}$$

The second inequality is obtained by applying Fubini's theorem. These terms are finite since

$$\int_{\mathbb{R}} |f(x)|^{3/2}\,dx \le \|f\|_\infty^{1/2} \int_{\mathbb{R}} f(x)\,dx = \|f\|_\infty^{1/2} < +\infty.$$

Thus, with $p = 3/2 \in]1,2]$, the function $(x,y) \longmapsto f^p(x+y)f^p(y)|x|^{1-p}$ is integrable. □

For instance, if ρ has a bounded support and a density which is even and continuous on it, then the hypothesis of the theorem are fulfilled.

We end this section by computing the law of $(S_n/n, T_n/n)$ under $\widetilde{\mu}_{n,\rho}$ and explaining the strategy for proving these results.

We denote by $\widetilde{\nu}_{n,\rho}$ the law of $(S_n/n, T_n/n)$ under $\rho^{\otimes n}$. We have

$$\forall (x_1,\dots,x_n) \in \mathbb{R}^n \qquad \frac{(x_1+\cdots+x_n)^2}{x_1^2+\cdots+x_n^2} = n\frac{((x_1+\cdots+x_n)/n)^2}{(x_1^2+\cdots+x_n^2)/n}.$$

Hence, for any bounded measurable function $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$,

$$\mathbb{E}_{\widetilde{\mu}_{n,\rho}}\left(f\left(\frac{S_n}{n},\frac{T_n}{n}\right)\right) = \frac{1}{Z_n}\int_{\mathbb{R}^2} f(x,y)\exp\left(\frac{nx^2}{2y}\right)\mathbb{1}_{\{y>0\}}\,d\widetilde{\nu}_{n,\rho}(x,y).$$

By convexity of $t \longmapsto t^2$, we have $S_n^2 \le nT_n$ for any $n \ge 1$. We define

$$\Delta = \{\,(x,y)\in\mathbb{R}^2 : x^2 \le y\,\} \qquad \text{and} \qquad \Delta^* = \Delta\backslash\{(0,0)\}.$$

Thus $\widetilde{\nu}_{n,\rho}(\Delta^c) = 0$. Therefore we have the following proposition:

Proposition 4. *Under $\widetilde{\mu}_{n,\rho}$, the law of $(S_n/n, T_n/n)$ is*

$$\frac{\exp\left(\dfrac{nx^2}{2y}\right)\mathbb{1}_{\Delta^*}(x,y)\,d\widetilde{\nu}_{n,\rho}(x,y)}{\displaystyle\int_{\Delta^*}\exp\left(\frac{ns^2}{2t}\right)\,d\widetilde{\nu}_{n,\rho}(s,t)}.$$

We denote by ν_ρ the law of (Z, Z^2) where Z is a random variable with distribution ρ. The Log-Laplace Λ of ν_ρ is the map defined on $\mathbb{R}^2$ by

$$\forall (u,v)\in\mathbb{R}^2 \qquad \Lambda(u,v) = \ln\int_{\mathbb{R}^2} e^{us+vt}\,d\nu_\rho(s,t) = \ln\int_{\mathbb{R}} e^{uz+vz^2}\,d\rho(z),$$

and the Cramér transform I of ν_ρ is defined on $\mathbb{R}^2$ by

$$\forall (x,y)\in\mathbb{R}^2 \qquad I(x,y) = \sup_{(u,v)\in\mathbb{R}^2}(xu+yv-\Lambda(u,v)).$$

For $n \ge 1$, under $\rho^{\otimes n}$, $(S_n/n, T_n/n)$ is the sum of n independent and identically distributed random variables with distribution ν_ρ. We refer to the appendix B for some definitions and results on large deviations, especially Cramér's theorem (theorem B.4) which states that, if Λ is finite in the neighbourhood of $(0,0)$, then

I is a good rate function and $(\widetilde{\nu}_{n,\rho})_{n\geq 1}$ satisfies the large deviations principle with speed n, governed by I.

Here is a classical heuristics on large deviations, suggested by a consequence of Varadhan's lemma (see theorem II.7.2 of [13]) : as n goes to $+\infty$, the law of $(S_n/n, T_n/n)$ under $\widetilde{\mu}_{n,\rho}$ concentrates exponentially fast on the minima on Δ^* of the function

$$G = I - F - \inf_{\Delta^*}(I - F),$$

where F is the map defined by

$$\forall (x,y) \in \mathbb{R}\times\mathbb{R}\backslash\{0\} \qquad F(x,y) = \frac{x^2}{2y}.$$

If G has a unique minimum at $(x_0, y_0) \in \Delta^*$, then, under $\widetilde{\mu}_{n,\rho}$, $(S_n/n, T_n/n)$ converges in probability to (x_0, y_0). Moreover, the large deviations principle suggests that, for n large enough, $\widetilde{\nu}_{n,\rho}$ can roughly be approximated by the distribution $C_n \exp(-nI(x,y))\,dx\,dy$ where C_n is a normalizing constant. Thus, for each bounded continuous function h and $\alpha, \beta > 0$,

$$\mathbb{E}_{\tilde{\mu}_n}\left(h\left(\frac{S_n - nx_0}{n^{1-\alpha}}\right)\right) \approx \frac{\displaystyle\int_{\Delta^*} h((x-x_0)n^{\alpha})\exp\left(-nG(x,y)\right)\,dx\,dy}{\displaystyle\int_{\Delta^*}\exp\left(-nG(x,y)\right)\,dx\,dy}$$

$$\approx \frac{\displaystyle\int_{\Delta^*} h(x)\exp\left(-nG\left(xn^{-\alpha}+x_0, yn^{-\beta}+y_0\right)\right)\,dx\,dy}{\displaystyle\int_{\Delta^*}\exp\left(-nG\left(xn^{-\alpha}+x_0, yn^{-\beta}+y_0\right)\right)\,dx\,dy}.$$

We use then Laplace's method. The key point is the study of the function G in the neighbourhood of its minimum (x_0, y_0). We find four positive values A, B, $a \in \mathbb{N}$ and $b \in \mathbb{N}$ such that, uniformly on a neighbourhood of (x_0, y_0),

$$-nG\left(xn^{-1/a}+x_0, yn^{-1/b}+y_0\right) \underset{n\to\infty}{\longrightarrow} -Ax^a - By^b.$$

We prove that $I-F$ has a unique minimum at $(0,\sigma^2)$ on Δ^* in section 4. Next we give the proof of theorem 1 in section 5, with the help of a variant of Varadhan's lemma. Finally we compute the expansion of $I-F$ around $(0,\sigma^2)$ in section 6 and we prove theorem 2 with Laplace's method in section 7. Throughout these proofs we use some general results on the Cramér transform, stated in appendix A.

4 Minimum of $I - F$ on Δ^*

Let ρ be a symmetric probability measure on $\mathbb{R}$. In this section, we will use proposition A.4 in appendix to show an inequality between I and F.

We denote by ν_ρ the distribution of (Z, Z^2) when Z is a random variable with law ρ. If the support of ρ contains at least three points then ν_ρ is a non-degenerate measure on $\mathbb{R}^2$ (see the first paragraphs of appendix A). We denote by $\mathcal{C}$ the convex hull of the set $\{\,(x,x^2) : x \text{ is in the support of } \rho\,\}$. The function

$$\Lambda : (u,v) \in \mathbb{R}^2 \longmapsto \ln\int_{\mathbb{R}} e^{uz+vz^2}\,d\rho(z)$$

is the Log-Laplace of ν_ρ and its domain of definition D_Λ contains $\mathbb{R}\times]-\infty,0[$, thus its interior is non-empty. Let I be the Cramér transform of ν_ρ. We denote by D_I its domain of definition and by $A_I = \nabla\Lambda(\overset{o}{D}_\Lambda)$ its admissible domain (see definition A.3 in appendix).

Using Jensen's inequality, we get that $I(0,\sigma^2) = 0$. Moreover the infimum of $I - F$ on Δ^* belongs to $[-1/2, 0]$. The function I is even in the first variable. Indeed, if $(x, y) \in \mathbb{R}^2$, then

$$\begin{aligned} I(-x,y) &= \sup_{(u,v)\in\mathbb{R}^2} \left(-xu + yv - \ln \int_{\mathbb{R}} e^{uz+vz^2}\, d\rho(z) \right) \\ &= \sup_{(u,v)\in\mathbb{R}^2} \left(xu + yv - \ln \int_{\mathbb{R}} e^{-uz+vz^2}\, d\rho(z) \right) = I(x,y). \end{aligned}$$

Assume that $I-F$ has a unique minimum (x_0, y_0) on Δ^*. Then $(-x_0, y_0)$ is also a minimum of $I - F$. The uniqueness of the minimum implies that $x_0 = 0$ so that $I - F$ is non-negative on Δ^*. Finally, since $I(0,\sigma^2) = 0$, we have $y_0 = \sigma^2$.

Consider first the case of a Bernoulli distribution for which ν_ρ is degenerate. Let $c > 0$. Suppose that $\rho = (\delta_{-c} + \delta_c)/2$. The law ρ is centered and its variance is c^2. We can compute Λ and I explicitly :

$$\forall (u,v) \in \mathbb{R}^2 \qquad \Lambda(u,v) = vc^2 + \ln\cosh(uc).$$

For any $(x,y) \notin [-c,c] \times \{c^2\}$, $I(x,y) = +\infty$ and

$$\forall x \in]-c,c[\qquad I(x,c^2) = \frac{1}{2c}\left((c+x)\ln(c+x) + (c-x)\ln(c-x)\right) - \ln c.$$

The study of the function $x \longmapsto I(x,c^2) - x^2/(2c^2)$ shows that, in the Bernoulli case, $I - F$ has a unique minimum at $(0,\sigma^2)$. More generally we have the following lemma:

Lemma 5. *Let $c > 0$. We define*

$$\phi_c : x \in \mathbb{R} \longmapsto \sup_{u\in\mathbb{R}} (\, ux - \ln\cosh(uc)\,).$$

The function $x \longmapsto \phi_c(x) - x^2/(2c^2)$ is increasing on $[0,c]$, decreasing on $[-c,0]$ and null in 0.

Notice that the Bernoulli case is special since, if X is a random variable with distribution $\rho = (\delta_{-c} + \delta_c)/2$, then $X^2 = c^2$ almost surely. Thus

$$\begin{aligned} &\frac{1}{Z_n} \exp\left(\frac{1}{2} \frac{(x_1 + \cdots + x_n)^2}{x_1^2 + \cdots + x_n^2} \right) \mathbb{1}_{\{x_1^2+\cdots+x_n^2>0\}} \prod_{i=1}^n d\rho(x_i) \\ &\qquad\qquad = \frac{1}{Z_n} \exp\left(\frac{(x_1 + \cdots + x_n)^2}{2nc^2} \right) \prod_{i=1}^n d\rho(x_i). \end{aligned}$$

This is exactly the classical Curie-Weiss model at the critical point.

In the following, we suppose that the support of ν_ρ contains at least three distinct points. We first show that, if D_Λ is an open subset of $\mathbb{R}^2$, then $I - F$ has a unique minimum at $(0,\sigma^2)$. To this end, we use proposition A.4 in appendix which states that I is differentiable on $A_I = \overset{o}{D}_I = \overset{o}{\mathcal{C}}$. Moreover, if $(x,y) \longmapsto (u(x,y), v(x,y))$ is the inverse function of $\nabla\Lambda$, then

$$\forall (x,y) \in \overset{o}{D}_I \qquad \frac{\partial I}{\partial x}(x,y) = u(x,y).$$

If we show that $u(x,y) > x/y$ for $x, y > 0$, then, by integrating this inequality,

$$\forall (x,y) \in \overset{o}{D}_I \qquad 0 \le \varepsilon < x \implies I(x,y) - \frac{x^2}{2y} > I(\varepsilon, y) - \frac{\varepsilon^2}{2y}.$$

To obtain that $I - F$ has a unique minimum at $(0,\sigma^2)$, it is enough to extend this inequality to the boundary points of D_I (if they exist). We conclude by using the fact that I is even in its first variable.

The following lemma is the key result to establish the uniqueness of the minimum of $I - F$, when ρ is symmetric.

Lemma 6. *Let ρ be a symmetric probability measure whose support contains at least three points. For $(x,y) \in A_I$, we have $u(x,y) = 0$ if $x = 0$ and*

$$u(x,y) > \frac{x}{y} \quad \text{if} \quad x > 0,$$
$$u(x,y) < \frac{x}{y} \quad \text{if} \quad x < 0.$$

Proof. The vector $(u,v) = (u(x,y), v(x,y))$ verifies

$$(x,y) = \nabla\Lambda(u,v) = \left(\frac{\displaystyle\int_{\mathbb{R}} z e^{uz+vz^2}\, d\rho(z)}{\displaystyle\int_{\mathbb{R}} e^{uz+vz^2}\, d\rho(z)}, \frac{\displaystyle\int_{\mathbb{R}} z^2 e^{uz+vz^2}\, d\rho(z)}{\displaystyle\int_{\mathbb{R}} e^{uz+vz^2}\, d\rho(z)} \right).$$

The distribution ρ is symmetric, thus

$$\int_{\mathbb{R}} z e^{uz+vz^2}\, d\rho(z) = \int_0^{+\infty} 2z \sinh(uz) e^{vz^2}\, d\rho(z).$$

This formula shows that u and x have the same sign. Moreover for any $z \ge 0$, $\tanh(z) \le z$ thus, if $x > 0$ then $\sinh(uz) \le uz\cosh(uz)$. The equality holds if and only if $uz = 0$. Therefore, using the symmetry of ρ,

$$x < u \frac{\displaystyle\int_0^{+\infty} 2z^2 \cosh(uz) e^{vz^2}\, d\rho(z)}{\displaystyle\int_{\mathbb{R}} e^{uz+vz^2}\, d\rho(z)} = u \frac{\displaystyle\int_{\mathbb{R}} z^2 e^{uz+vz^2}\, d\rho(z)}{\displaystyle\int_{\mathbb{R}} e^{uz+vz^2}\, d\rho(z)} = uy.$$

Since $x > 0$, $u > 0$ and $y > 0$, we conclude that $u > x/y$. Similarly, we show that if $x < 0$ then $u < x/y$. □

We can now prove the following inequality:

Proposition 7. *If ρ is a symmetric probability measure on $\mathbb{R}$ with variance $\sigma^2 > 0$ and such that D_Λ is an open subset of $\mathbb{R}^2$ then*

$$\forall (x,\varepsilon,y) \in \mathbb{R}\times\mathbb{R}\times\mathbb{R}\backslash\{0\} \qquad 0 \leq \varepsilon < x \implies I(x,y) - \frac{x^2}{2y} \geq I(\varepsilon,y) - \frac{\varepsilon^2}{2y}.$$

This inequality is strict if $(\varepsilon, y) \in \overset{o}{D}_I$.

The inequality is also true for $x < \varepsilon \leq 0$ since I is even in its first variable. In corollary 12, we shall extend the inequality to any symmetric distribution on $\mathbb{R}$.

Proof. We have already treated the Bernoulli case. We assume next that the support of ρ contains at least three points. The Cramér transform I is $\mathcal{C}^\infty$ on $\overset{o}{D}_I$ and

$$\forall (x,y) \in \overset{o}{D}_I \qquad \frac{\partial I}{\partial x}(x,y) = u(x,y).$$

Let us examine the structure of the set D_I. We put

$$\forall y > 0 \qquad D_{I,y} = \{\, x \in \mathbb{R} : (x,y) \in D_I \,\}.$$

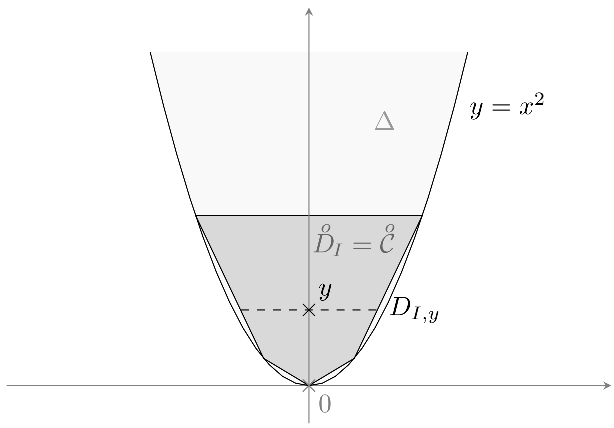
Let $y > 0$ be such that $(x,y) \in \overset{o}{D}_I$ for some $x \in \mathbb{R}$. The set $D_{I,y}$ is a convex subset of $\mathbb{R}$. Moreover $x \longmapsto I(x,y)$ is even, therefore $\overset{o}{D}_{I,y}$ (the interior of $D_{I,y}$ as a subset of $\mathbb{R}$) is an open interval $]-a(y), a(y)[$ with $a(y) \in [0, \sqrt{y}]$. Lemma 6 implies that $u(t,y) > t/y$ for any $t \in]0, a(y)[$. Thus, for any $x \in \overset{o}{D}_{I,y} \cap]0,+\infty[$,

$$\forall \varepsilon \in [0,x[\qquad I(x,y) - I(\varepsilon,y) = \int_\varepsilon^x u(t,y)\,dt > \int_\varepsilon^x \frac{t}{y}\,dt = \frac{x^2}{2y} - \frac{\varepsilon^2}{2y}.$$

There is no problem of definition at $y = 0$ since $\overset{o}{D}_I \subset \Delta^*$ does not contain $\mathbb{R}\times\{0\}$ and $\overset{o}{D}_{I,0} = \varnothing$. Moreover

$$x \longmapsto \frac{I(x,y) - I(\varepsilon,y)}{x - \varepsilon}$$

is non-decreasing on $D_{I,y}\backslash\{\varepsilon\}$ since I is convex. Therefore, if $-a(y)$ and $a(y)$ belong to $D_{I,y}$, then the previous inequality extends to $x = -a(y)$ and $x = a(y)$.



Case where ρ is symmetric discrete and charges 5 points

We have shown that

$$\forall (x,y) \in D_I \qquad y>0,\, 0\le \varepsilon < x \implies I(x,y) - I(\varepsilon,y) > \frac{x^2}{2y} - \frac{\varepsilon^2}{2y},$$

except for the points (x,y) of the superior and inferior borders of D_I, if they exist. More precisely, we set

$$K^2 = \inf \{\, x^2 : x \text{ is in the support of } \rho \} \geq 0$$

and

$$L^2 = \sup \{\, x^2 : x \text{ is in the support of } \rho \} \leq +\infty.$$

If $K=0$ and $L=+\infty$ then the inequality is already proved on $D_I \backslash \{(0,0)\}$. Suppose that $K^2>0$. Let $y=K^2$ and $x\in\mathbb{R}$. We define

$$f : (u,v) \in \mathbb{R}^2 \longmapsto ux + vK^2 - \Lambda(u,v).$$

Denoting $c_K = \rho(\{K\})$, we have for all $(u,v)\in\mathbb{R}^2$,

$$f(u,v) = ux - \ln(2c_K \cosh(uK)) - \ln \int_{\mathbb{R}\backslash[-K,K]} e^{uz+v(z^2-K^2)}\, d\rho(z).$$

For any $z\in \mathbb{R}\backslash]-K,K[$, the function $v\longmapsto \exp(v(z^2-K^2))$ is non-decreasing. Therefore

$$\begin{aligned}\sup_{v\in\mathbb{R}} f(u,v) &= ux - \ln(2c_K\cosh(uK)) - \ln\left(\lim_{v\to-\infty}\int_{\mathbb{R}\backslash[-K,K]} e^{uz+v(z^2-K^2)}\, d\rho(z)\right)\\ &= ux - \ln(2c_K\cosh(uK)),\end{aligned}$$

by the dominated convergence theorem. Indeed

$$\forall z\in\mathbb{R}\backslash[-K,K] \quad \forall v<-1 \qquad \left| e^{uz+v(z^2-K^2)} \right| \leq e^{uz-(z^2-K^2)}$$

and the map $z\in\mathbb{R}\backslash[-K,K] \longmapsto e^{uz-(z^2-K^2)}$ is integrable with respect to ρ since it is bounded (it is continuous and goes to 0 when $|z|$ goes to $+\infty$). Hence

$$I(x,K^2) = \sup_{u,v\in\mathbb{R}} f(u,v) = \sup_{u\in\mathbb{R}} \{\, ux - \ln(2c_K\cosh(uK)) \,\}.$$

In fact, we come back to the Bernoulli case. The reason is that, if we condition on $T_n = K^2$ in our model, then for any i, $X_n^i = -K$ or K.

If $c_K=0$ then for any $x\neq 0$, $I(x,K^2)=+\infty$ so that the (large) inequality is verified for $y=K^2$. If $c_K>0$, then lemma 5 implies that, for all $(x,y)\in\mathbb{R}^2$,

$$0\le\varepsilon<x\le K \implies I(x,K^2) - I(\varepsilon,K^2) = \phi_K(x) - \phi_K(\varepsilon) > \frac{x^2}{2K^2} - \frac{\varepsilon^2}{2K^2}.$$

If $L<+\infty$ then we show similarly the inequality for $y=L^2$. Therefore

$$\forall (x,y)\in D_I\backslash\{(0,0)\} \qquad 0\le\varepsilon<x \implies I(x,y) - \frac{x^2}{2y} \geq I(\varepsilon,y) - \frac{\varepsilon^2}{2y}$$

and this inequality is strict if $(\varepsilon,y)\in \overset{o}{D}_I$. Finally we notice that for any $y\in\mathbb{R}$, by the convexity and the symmetry of $x\longmapsto I(x,y)$, if $I(\varepsilon,y)=+\infty$ then for all $x > \varepsilon$, $I(x,y) = +\infty$. Therefore the inequality extends to each subset of $\mathbb{R}^2$ which does not contain $\mathbb{R} \times \{0\}$. □

From the arguments in the previous proof, we notice that if we take $x = y = 0$, then for any $u \in \mathbb{R}$, the function $v \longmapsto \Lambda(u,v)$ is non-decreasing on $\mathbb{R}$. Therefore

$$\inf_{v\in\mathbb{R}} \Lambda(u,v) = \lim_{v\to-\infty} \Lambda(u,v) = \lim_{v\to-\infty} \left(\ln \rho(\{0\}) + \ln \int_{\mathbb{R}\setminus\{0\}} e^{uz+vz^2}\, d\rho(z) \right).$$

By the dominated convergence theorem, the last integral is equal to $\ln \rho(\{0\})$. Hence

$$\inf_{u,v\in\mathbb{R}^2} \Lambda(u,v) = \inf_{u\in\mathbb{R}} (\ln \rho(\{0\})) = \ln \rho(\{0\}).$$

This is valid for any probability measure ρ on $\mathbb{R}$. This yields the following lemma:

Lemma 8. *If ρ is a probability measure on $\mathbb{R}$ then $I(0,0) = -\ln \rho(\{0\})$.*

A consequence of proposition 7 and the fact that I is even in its first variable is that, if D_Λ is an open subset of $\mathbb{R}^2$, then the function $I - F$ has a unique minimum on Δ^* at $(0,\sigma^2)$. Now we will extend this result to any symmetric probability measure such that $(0,0) \in \overset{o}{D}_\Lambda$. For this we need Mosco's theorem, which we restate next.

Definition 9. *Let f and f_n, $n \in \mathbb{N}$, be convex functions from $\mathbb{R}^d$ to $[-\infty,+\infty]$. The sequence $(f_n)_{n\in\mathbb{N}}$ is said to Mosco converge to f if for any $x \in \mathbb{R}^d$,*
⋆ for each sequence $(x_n)_{n\in\mathbb{N}}$ in $\mathbb{R}^d$ converging to x,

$$\liminf_{n\to+\infty} f_n(x_n) \geq f(x),$$

⋆ there exists a sequence $(x_n)_{n\in\mathbb{N}}$ in $\mathbb{R}^d$ converging to x and such that

$$\limsup_{n\to+\infty} f_n(x_n) \leq f(x).$$

If f is a convex function from $\mathbb{R}^d$ to $[-\infty,+\infty]$, we denote by f^* its Fenchel-Legendre transform f^*. We have the following theorem (see [18] for a proof):

Theorem 10 (Mosco)**.** *Let f and f_n, $n \in \mathbb{N}$, be functions from $\mathbb{R}^d$ to $[-\infty,+\infty]$ which are convex and lower semi-continuous. Then $(f_n)_{n\in\mathbb{N}}$ Mosco converges to f if and only if $(f_n^*)_{n\in\mathbb{N}}$ Mosco converges to f^*.*

Proposition 11. *Let ν be a probability measure on $\mathbb{R}^d$. We denote by L its Log-Laplace. Let $(K_n)_{n\in\mathbb{N}}$ be a non-decreasing sequence of compact sets whose union is $\mathbb{R}^d$. For all $n \in \mathbb{N}$, we set $\nu_n = \nu(.|K_n)$ the probability ν conditioned by K_n and we denote by L_n its Log-Laplace. Then $(L_n)_{n\in\mathbb{N}}$ Mosco converges to L.*

Proof. For n large enough, the compact set K_n meets the support of ν. Thus, for n large enough and $\lambda \in \mathbb{R}^d$, we have

$$L_n(\lambda) = \ln \int_{\mathbb{R}^d} e^{\langle\lambda,z\rangle}\, d\nu_n(z) = \ln \int_{K_n} e^{\langle\lambda,z\rangle}\, d\nu(z) - \ln \nu(K_n).$$

By the monotone convergence theorem,

$$\lim_{n\to+\infty} L_n(\lambda) = \ln \int_{\mathbb{R}^d} \lim_{n\to+\infty} \left(\mathbb{1}_{K_n}(z) e^{\langle\lambda,z\rangle} \right) d\nu(z) - \lim_{n\to+\infty} \ln \nu(K_n) = L(\lambda).$$

Hence the second condition of Mosco convergence (with the limsup) is satisfied with the sequence $(\lambda_n)_{n\in\mathbb{N}}$ constant equal to λ.

Let $\lambda \in \mathbb{R}^d$ and $(\lambda_n)_{n\in\mathbb{N}}$ be any sequence converging to λ. Fatou's lemma implies that

$$\exp L(\lambda) = \int_{\mathbb{R}^d} \liminf_{n\to+\infty} \mathbb{1}_{K_n}(z)e^{\langle\lambda_n,z\rangle}\,d\nu(z) \le \liminf_{n\to+\infty} \int_{\mathbb{R}^d} \mathbb{1}_{K_n}(z)e^{\langle\lambda_n,z\rangle}\,d\nu(z).$$

Therefore

$$L(\lambda) \le \liminf_{n\to+\infty} \left(L_n(\lambda_n) + \ln\nu(K_n)\right) = \liminf_{n\to+\infty} L_n(\lambda_n).$$

Thus the first condition of Mosco convergence (with the liminf) is verified and the proposition is proved. □

Corollary 12. *If ρ is a symmetric and non-degenerate probability measure on $\mathbb{R}$ then*

$$\forall (x,y)\in\Delta^* \quad \forall\varepsilon\in[0,|x|[\qquad I(x,y)-\frac{x^2}{2y} \ge I(\varepsilon,y)-\frac{\varepsilon^2}{2y}.$$

Proof. For any $n\in\mathbb{N}$, we put $K_n=[-n,n]^2$. For n large enough so that K_n meets the support of ν_ρ, we define $\nu_n=\nu_\rho(.|K_n)$, Λ_n its Log-Laplace and I_n its Fenchel-Legendre transform. For all $(u,v)\in\mathbb{R}^2$,

$$\Lambda_n(u,v) = \ln\int_{K_n} e^{us+vt}\,d\nu_\rho(s,t) - \ln\nu_\rho(K_n) \le \Lambda(u,v) - \ln\nu_\rho(K_n).$$

Applying the Fenchel-Legendre transformation, we get

$$\forall(\varepsilon,y)\in\mathbb{R}^2 \qquad I(\varepsilon,y)\le I_n(\varepsilon,y)-\ln\nu_\rho(K_n).$$

Moreover the measure ν_n has a bounded support thus proposition 7 and the previous inequality imply that, for any $(x,\varepsilon,y)\in\mathbb{R}\times\mathbb{R}\times]0,+\infty[$,

$$0\le\varepsilon<x \quad\Longrightarrow\quad I(\varepsilon,y)-\frac{\varepsilon^2}{2y}+\frac{x^2}{2y}\le I_n(x,y)-\ln\nu_\rho(K_n).$$

It follows from proposition 11 that $(\Lambda_n)_{n\in\mathbb{N}}$ Mosco converges to Λ. Hence, by Mosco's theorem, $(I_n)_{n\in\mathbb{N}}$ Mosco converges to I. In particular, for $(x,y)\in\mathbb{R}^2$ such that $y>0$ and $x>\varepsilon$, there exists a sequence $(x_n,y_n)\in\mathbb{R}^2$ converging to (x,y) and such that

$$\limsup_{n\to+\infty} I_n(x_n,y_n)\le I(x,y).$$

Since $y>0$ and $x>\varepsilon$, there exists $n_0\ge1$ such that $y_n>0$ and $x_n>\varepsilon$ for all $n\ge n_0$. Therefore

$$\forall n\ge n_0 \qquad I(\varepsilon,y_n)-\frac{\varepsilon^2}{2y_n}+\frac{x_n^2}{2y_n}\le I_n(x_n,y_n)-\ln\nu_\rho(K_n).$$

Moreover $\nu_\rho(K_n)\to1$ when $n\to\infty$. Hence

$$\limsup_{n\to+\infty} I(\varepsilon,y_n)-\frac{\varepsilon^2}{2y}+\frac{x^2}{2y}\le I(x,y).$$

Finally I is lower semi-continuous, thus

$$\liminf_{n\to+\infty} I(\varepsilon, y_n) \geq I(\varepsilon, y).$$

This implies the announced inequality. □

We can now show that $I - F$ has a unique minimum on Δ^* at $(0, \sigma^2)$:

Proposition 13. *If ρ is a symmetric probability measure on $\mathbb{R}$ with variance $\sigma^2 > 0$ and such that Λ is finite in a neighbourhood of $(0,0)$ then*

$$(x, y) \in \Delta^* \longmapsto I(x, y) - \frac{x^2}{2y}$$

has a unique minimum at $(0, \sigma^2)$ where it is equal to 0.

Proof. Corollary 12 implies that

$$\forall (x, y) \in \Delta^* \qquad I(x, y) - \frac{x^2}{2y} \geq I(0, y).$$

Therefore $I - F$ is a non-negative function. Since $(0,0) \in \overset{o}{D}_\Lambda$, the function $I(0, .)$ has a unique minimum at σ^2 (see theorems 25.1 and 27.1 of [21]). As a consequence, if $I - F$ has a minimum on Δ^* at (x_0, y_0), then $y_0 = \sigma^2$ and $I(x_0, \sigma^2) = x_0^2/(2\sigma^2)$.

Moreover $(0, \sigma^2) \in A_I$ thus there exists $\varepsilon > 0$ such that B_ε, the open ball of radius ε centered at $(0, \sigma^2)$, is included in A_I. If (x, y) realizes a minimum of $I - F$ on B_ε then

$$(u(x, y), v(x, y)) = \nabla I(x, y) = \nabla F(x, y) = \Big(x/y, -x^2/(2y^2)\Big).$$

It follows from lemma 6 that $x = 0$ and thus $u(x, y) = v(x, y) = 0$. Therefore $(x, y) = (0, \sigma^2)$. Hence

$$\forall x \in]-\varepsilon, 0[\cap]0, \varepsilon[\qquad I(x, \sigma^2) - \frac{x^2}{2\sigma^2} > 0.$$

Applying corollary 12 with $\varepsilon/2$, we see that the above inequality holds for any $x \neq 0$. It follows that $x_0 = 0$. □

5 Proof of theorem 1 with a variant of Varadhan's lemma

Let ρ be a symmetric probability measure on $\mathbb{R}$ with positive variance σ^2 and such that $(0,0) \in \overset{o}{D}_\Lambda$. The heuristics at the end of section 3 and proposition 13 suggest that, as n goes to $+\infty$, the law of $(S_n/n, T_n/n)$ under $\widetilde{\mu}_{n,\rho}$ concentrates exponentially fast on $(0, \sigma^2)$, the minimum of $I - F$.

Yet, in spite of the expression given in proposition 4, we cannot apply Varadhan's lemma (theorem II.7.2 of [13]) directly since Δ^* is not a closed set and F is not continuous on Δ.

In subsection 5.a), we prove a variant of Varadhan's lemma. We give the proof of theorem 1 in subsection 5.b).

a) Around Varadhan's lemma

Proposition 14. *Let ρ be a probability measure on $\mathbb{R}$. We denote by $\widetilde{\nu}_{n,\rho}$ the distribution of $(S_n/n, T_n/n)$ under $\rho^{\otimes n}$. We have*

$$\liminf_{n\to+\infty} \frac{1}{n} \ln \int_{\Delta^*} \exp\left(\frac{nx^2}{2y}\right) d\widetilde{\nu}_{n,\rho}(x,y) \geq 0.$$

Suppose that ρ is non-degenerate, symmetric and that $(0,0) \in \overset{o}{D}_\Lambda$. We assume that there exists $r > 0$ such that $M_r + \ln \rho(\{0\}) < 0$ with

$$M_r = \sup\left\{ \frac{x^2}{2y} : (x,y) \in \mathcal{C} \cap \mathrm{B}_r \backslash \{(0,0)\} \right\},$$

where B_r is the open ball of radius r centered at $(0,0)$ and $\mathcal{C}$ is the closed convex hull of $\{ (x,x^2) : x$ is in the support of $\rho \}$. If A is a closed subset of $\mathbb{R}^2$ which does not contain $(0,\sigma^2)$ then

$$\limsup_{n\to+\infty} \frac{1}{n} \ln \int_{\Delta^* \cap A} \exp\left(\frac{nx^2}{2y}\right) d\widetilde{\nu}_{n,\rho}(x,y) < 0.$$

Let us give first some sufficient conditions to fulfill the hypothesis of the proposition. To ensure that there exists $r > 0$ such that $M_r + \ln \rho(\{0\}) < 0$, it is enough that one of the following conditions is satisfied:

(a) ρ has a density.

(b) $\rho(\{0\}) < 1/\sqrt{e}$.

(c) There exists $c > 0$ such that $\rho(]0,c[) = 0$.

(d) ρ is the sum of a finite number of Dirac masses.

Indeed, the function F is bounded by $1/2$ on $\mathcal{C}\backslash\{(0,0)\} \subset \Delta^*$, thus for any $r > 0$, $M_r \leq 1/2$. Therefore, if ρ has a density, or more generally if $\rho(\{0\}) < e^{-1/2}$, then for all $r > 0$, $M_r + \ln \rho(\{0\}) < 0$.

On the other hand, if there exists $c > 0$ such that $]0,c[$ does not intersect the support of ρ (especially if ρ is the sum of a finite number of Dirac masses) then

$$\mathcal{C} \subset \{ (x,y) \in \mathbb{R}^2 : c|x| \leq y \}.$$

Therefore

$$\forall (x,y) \in \mathcal{C} \cap \mathrm{B}_r \backslash \{(0,0)\} \qquad \frac{x^2}{2y} = \frac{c|x|^2}{2cy} \leq \frac{|x|}{2c} \leq \frac{r}{2c}.$$

Hence for any $r > 0$, $M_r < r/2c$. Since ρ is non-degenerate, $\rho(\{0\}) < 1$, thus there exists $r > 0$ such that $\ln \rho(\{0\}) + r/2c < 0$. Therefore the conditions (c) and (d) imply that $M_r + \ln \rho(\{0\}) < 0$.

Proof of proposition 14. The large deviations principle satisfied by $(\widetilde{\nu}_{n,\rho})_{n\geq 1}$ implies that

$$\liminf_{n\to+\infty} \frac{1}{n} \ln \int_{\Delta^*} \exp\left(\frac{nx^2}{2y}\right) d\widetilde{\nu}_{n,\rho}(x,y)$$
$$\geq \liminf_{n\to+\infty} \frac{1}{n} \ln \widetilde{\nu}_{n,\rho}(\Delta^*) \geq -\inf\left\{ I(x,y) : (x,y) \in \overset{o}{\Delta} \right\} = 0.$$

We prove now the second inequality. Let $\alpha > 0$. The function I is lower semi-continuous on $\mathbb{R}^2$, thus there exists a neighbourhood $\mathcal{U}$ of $(0,0)$ such that

$$\forall (x,y) \in \overline{\mathcal{U}} \qquad I(x,y) \geq (I(0,0) - \alpha) \wedge \frac{1}{\alpha} = (-\ln \rho(\{0\}) - \alpha) \wedge \frac{1}{\alpha}.$$

The above equality follows from lemma 8. By hypothesis, there exists $r > 0$ such that $M_r + \ln \rho(\{0\}) < 0$ thus, by choosing α sufficiently small, we can assume that

$$M_r + \ln \rho(\{0\}) + \alpha < 0 \qquad \text{and} \qquad M_r - \frac{1}{\alpha} < 0.$$

Since M_r decreases with r, we can take r small enough so that $\mathrm{B}_r \subset \mathcal{U}$. Notice next that $(S_n/n, T_n/n) \in \mathcal{C}$ almost surely. Therefore, setting $\mathcal{C}^* = \mathcal{C} \backslash \{(0,0)\}$,

$$\int_{\Delta^* \cap A} \exp\left(\frac{nx^2}{2y}\right) \, d\widetilde{\nu}_{n,\rho}(x,y) = \int_{\mathcal{C}^* \cap A} \exp\left(\frac{nx^2}{2y}\right) \, d\widetilde{\nu}_{n,\rho}(x,y).$$

Let us decompose

$$\mathcal{C}^* \cap A \subset (\mathcal{C}^* \cap \mathrm{B}_r) \cup (\mathcal{C} \cap \mathrm{B}_r^c \cap A).$$

We have

$$\int_{\mathcal{C}^* \cap \mathrm{B}_r} \exp\left(\frac{nx^2}{2y}\right) \, d\widetilde{\nu}_{n,\rho}(x,y) \leq \exp(nM_r) \, \widetilde{\nu}_{n,\rho}\,(\mathcal{U}).$$

The large deviation principle satisfied by $(\widetilde{\nu}_{n,\rho})_{n \geq 1}$ implies that

$$\begin{aligned}\limsup_{n \to +\infty} \frac{1}{n} \ln \int_{\mathcal{C}^* \cap \mathrm{B}_r} \exp\left(\frac{nx^2}{2y}\right) \, d\widetilde{\nu}_{n,\rho}(x,y) &\leq M_r - \inf_{\overline{\mathcal{U}}} I \\ &\leq (M_r + \ln \rho(\{0\}) + \alpha) \vee \left(M_r - \frac{1}{\alpha}\right).\end{aligned}$$

Next, the set $\mathcal{C} \cap \mathrm{B}_r^c \cap A$ is closed and does not contain $(0,0)$ thus the function F is continuous on this set. Moreover F is bounded on $\mathcal{C}^*$. Hence lemma B.3 in appendix and lemma 1.2.15 of [10] imply that

$$\begin{aligned}&\limsup_{n \to +\infty} \frac{1}{n} \ln \int_{\mathcal{C}^* \cap A} \exp\left(\frac{nx^2}{2y}\right) \, d\widetilde{\nu}_{n,\rho}(x,y) \\ &\qquad \leq \max\left(M_r + \ln \rho(\{0\}) + \alpha \,,\, M_r - \frac{1}{\alpha} \,,\, \sup_{\mathcal{C} \cap \mathrm{B}_r^c \cap A} (F - I) \right).\end{aligned}$$

Since ρ is symmetric and $(0,0) \in \overset{o}{D}_\Lambda$, proposition 13 implies that $G = I - F$ has a unique minimum at $(0, \sigma^2)$ on Δ^*. Suppose that the infimum of G over $\mathcal{C} \cap \mathrm{B}_r^c \cap A$ is null. Then there exists a sequence $(x_k, y_k)_{k \in \mathbb{N}}$ in $\mathcal{C} \cap \mathrm{B}_r^c \cap A \subset \Delta^*$ such that

$$\lim_{k \to +\infty} G(x_k, y_k) = \inf_{\mathcal{C} \cap \mathrm{B}_r^c \cap A} G = 0.$$

For k large enough, $G(x_k, y_k) \leq 1/2$ thus $I(x_k, y_k) \leq 1$, i.e., (x_k, y_k) belongs to the compact set $\{\,(u,v) \in \mathbb{R}^2 : I(u,v) \leq 1\,\}$. Up to the extraction of a subsequence, we suppose that $(x_k, y_k)_{k \in \mathbb{N}}$ converges to some (x_0, y_0), which belongs to the closed subset $\mathcal{C} \cap \mathrm{B}_r^c \cap A$. Moreover G is lower semi-continuous, hence

$$0 = \liminf_{k \to +\infty} G(x_k, y_k) \geq G(x_0, y_0) \geq 0.$$

Therefore $G(x_0, y_0) = 0$ and thus $(x_0, y_0) = (0, \sigma^2) \in \mathcal{C} \cap \mathrm{B}_r^c \cap A$, which is absurd since A does not contain $(0, \sigma^2)$. Thus the infimum of G over $\mathcal{C} \cap \mathrm{B}_r^c \cap A$ is positive. Therefore

$$\max\left(M_r + \ln \rho(\{0\}) + \alpha \, , \, M_r - \frac{1}{\alpha} \, , \sup_{\mathcal{C} \cap \mathrm{B}_r^c \cap A} (F - I) \right) < 0.$$

This proves the second inequality. □

b) Proof of theorem 1

Let ρ be a symmetric probability measure on $\mathbb{R}$ with positive variance σ^2 and such that

$$\exists v_0 > 0 \qquad \int_{\mathbb{R}} e^{v_0 z^2} \, d\rho(z) < +\infty.$$

This implies that $\mathbb{R} \times]-\infty, v_0[\subset D_\Lambda$ and thus $(0,0) \in \overset{o}{D}_\Lambda$. We assume that one of the four conditions given in the paragraph below proposition 14 is satisfied.

We denote by $\theta_{n,\rho}$ the distribution of $(S_n/n, T_n/n)$ under $\widetilde{\mu}_{n,\rho}$. Let U be an open neighbourhood of $(0, \sigma^2)$ in $\mathbb{R}^2$. Propositions 4 and 14 imply that

$$\begin{aligned} \limsup_{n \to +\infty} \frac{1}{n} \ln \theta_{n,\rho}(U^c) = \limsup_{n \to +\infty} \frac{1}{n} \ln \int_{\Delta^* \cap U^c} \exp\left(\frac{nx^2}{2y}\right) \, d\widetilde{\nu}_{n,\rho}(x,y) \\ - \liminf_{n \to +\infty} \frac{1}{n} \ln \int_{\Delta^*} \exp\left(\frac{nx^2}{2y}\right) \, d\widetilde{\nu}_{n,\rho}(x,y) < 0. \end{aligned}$$

Hence there exist $\varepsilon > 0$ and $n_0 \in \mathbb{N}$ such that for any $n > n_0$,

$$\theta_{n,\rho}(U^c) \leq e^{-n\varepsilon} \underset{n \to \infty}{\longrightarrow} 0.$$

Thus, for each open neighbourhood U of $(0, \sigma^2)$,

$$\lim_{n \to +\infty} \widetilde{\mu}_{n,\rho}\left(\left(\frac{S_n}{n}, \frac{T_n}{n} \right) \in U^c \right) = 0.$$

This means that, under $\widetilde{\mu}_{n,\rho}$, $(S_n/n, T_n/n)$ converges in probability to $(0, \sigma^2)$. This ends the proof of theorem 1.

6 Expansion of $I - F$ around its minimum

In this section, which may be omitted on a first reading, we compute the expansion of the function $I - F$ around $(0, \sigma^2)$, its minimum over Δ^*. These computations are crucial because they explain why the fluctuations in theorem 2 are of order $n^{3/4}$ and they give us the term in the exponential in the limiting law.

If ρ is a symmetric probability measure whose support contains at least three points and if $(0,0) \in \overset{o}{D}_L$ then $(0,\sigma^2) = \nabla\Lambda(0,0) \in \nabla\Lambda(\overset{o}{D}_\Lambda) = A_I$, the admissible domain of I. Proposition A.4 in appendix implies that I is $\mathcal{C}^\infty$ in the neighbourhood of $(0,\sigma^2)$ and that

$$\nabla I(0,\sigma^2) = (u(0,\sigma^2), v(0,\sigma^2)) = (\nabla\Lambda)^{-1}(0,\sigma^2) = (0,0),$$

$$\mathrm{D}^2_{(0,\sigma^2)} I = \left(\mathrm{D}^2_{(0,0)}\Lambda\right)^{-1} = \begin{pmatrix} \sigma^2 & 0 \\ 0 & \mu_4 - \sigma^4 \end{pmatrix}^{-1} = \begin{pmatrix} 1/\sigma^2 & 0 \\ 0 & 1/(\mu_4 - \sigma^4) \end{pmatrix},$$

since $\mathrm{D}^2_{(0,0)}\Lambda$ is the covariance matrix of ν_ρ. Hence, up to the second order, the expansion of $I - F$ in the neighbourhood of $(0,\sigma^2)$ is

$$I(x,y) - F(x,y) = \frac{(y-\sigma^2)^2}{2(\mu_4 - \sigma^4)} + o(\|x, y - \sigma^2\|^2).$$

We need to push further the expansion of $I - F$.

Consider the case of the Gaussian $\mathcal{N}(0,\sigma^2)$. We can compute explicitly I:

$$\forall (x,y) \in \Delta^* \qquad I(x,y) = \frac{1}{2}\left(\frac{y}{\sigma^2} - 1 - \ln\left(\frac{y - x^2}{\sigma^2}\right)\right).$$

In the neighbourhood of $(0,\sigma^2)$, we have

$$I(x,y) - F(x,y) \sim \frac{x^4}{4\sigma^4} + \frac{(y-\sigma^2)^2}{4\sigma^2}.$$

In fact, we have a similar expansion in a more general case:

Proposition 15. *If ρ is a symmetric probability measure on $\mathbb{R}$ whose support contains at least three points and such that $(0,0) \in \overset{o}{D}_\Lambda$ then I is $\mathcal{C}^\infty$ in the neighbourhood of $(0,\sigma^2)$. If μ_4 denotes the fourth moment of ρ then, when (x,y) goes to $(0,\sigma^2)$,*

$$I(x,y) - \frac{x^2}{2y} \sim \frac{(y-\sigma^2)^2}{2(\mu_4 - \sigma^4)} + \frac{\mu_4 x^4}{12\sigma^8}.$$

Proof. If $(0,0) \in \overset{o}{D}_\Lambda$ then $(0,\sigma^2) = \nabla\Lambda(0,0) \in \nabla\Lambda(\overset{o}{D}_\Lambda) = A_I$ and proposition A.4 in appendix implies that the function I is $\mathcal{C}^\infty$ on A_I. Moreover, if we denote the inverse function of $\nabla\Lambda$ by $(x,y) \longmapsto (u(x,y), v(x,y))$, then, for all $(x,y) \in A_I$,

$$\nabla I(x,y) = (u(x,y), v(x,y)) \qquad \text{and} \qquad \mathrm{D}^2_{(x,y)} I = \left(\mathrm{D}^2_{(u(x,y),v(x,y))}\Lambda\right)^{-1}.$$

The hypothesis $(0,0) \in \overset{o}{D}_\Lambda$ also implies that ρ has finite moments of all order. The expansion of F to the fourth order in the neighbourhood of $(0,\sigma^2)$ is

$$F(x,y) = \frac{x^2}{2\sigma^2} - \frac{x^2(y-\sigma^2)}{2\sigma^4} + \frac{x^2(y-\sigma^2)^2}{2\sigma^6} + o(\|x, y-\sigma^2\|^4).$$

Therefore, in the neighbourhood of $(0,0)$,

$$I(x,h+\sigma^2)-F(x,h+\sigma^2)=\frac{h^2}{2(\mu_4-\sigma^4)}+a_{3,0}x^3+a_{2,1}x^2h+a_{1,2}xh^2+a_{0,3}h^3 \\ +a_{4,0}x^4+a_{3,1}x^3h+a_{2,2}x^2h^2+a_{1,3}xh^3+a_{0,4}h^4+o(\|x,h\|^4),$$

with, for any $(i,j)\in\mathbb{N}$ such that $i+j\in\{3,4\}$,

$$a_{i,j}=\frac{1}{i!j!}\frac{\partial^{i+j}I}{\partial x^i\partial y^j}(0,\sigma^2),$$

except for

$$a_{2,1}=\frac{1}{2}\left(\frac{\partial^3 I}{\partial x^2\partial y}(0,\sigma^2)+\frac{1}{\sigma^4}\right)\qquad\text{and}\qquad a_{2,2}=\frac{1}{4}\frac{\partial^4 I}{\partial x^2\partial y^2}(0,\sigma^2)-\frac{1}{2\sigma^6}.$$

If we prove that $a_{4,0}>0$ then the terms xh^2, h^3, x^3h, x^2h^2, xh^3 and h^4 are negligible compared to $a_{4,0}x^4+a_{0,2}h^2$ when (x,h) goes to $(0,0)$. Next, the symmetry of $I-F$ in the first variable implies that $a_{3,0}=0$. If we show that $a_{2,1}=0$ then, when $(x,y)\to(0,\sigma^2)$,

$$I(x,y)-F(x,y)=\left(\frac{(y-\sigma^2)^2}{2(\mu_4-\sigma^4)}+a_{4,0}x^4\right)(1+o(1)).$$

To conclude it is enough to show that $a_{2,1}=0$ and $a_{4,0}=\mu_4/(12\sigma^8)$, that is

$$\frac{\partial^3 I}{\partial x^2\partial y}(0,\sigma^2)=-\frac{1}{\sigma^4}\qquad\text{and}\qquad\frac{\partial^4 I}{\partial x^4}(0,\sigma^8)=\frac{2\mu_4}{\sigma^2}.$$

For any $j\in\mathbb{N}$, we introduce the function f_j defined on $\overset{o}{D}_\Lambda$ by

$$\forall(u,v)\in\overset{o}{D}_\Lambda\qquad f_j(u,v)=\int_{\mathbb{R}}x^je^{ux+vx^2}\,d\rho(x)\left(\int_{\mathbb{R}}e^{ux+vx^2}\,d\rho(x)\right)^{-1}.$$

These functions are $\mathcal{C}^\infty$ on $\overset{o}{D}_\Lambda$ and they verify the following properties:
⋆ f_0 is the identity function on $\mathbb{R}^2$ and

$$f_1=\frac{\partial\Lambda}{\partial u}\qquad\text{and}\qquad f_2=\frac{\partial\Lambda}{\partial v}.$$

⋆ For all $j\in\mathbb{N}$, $f_j(0,0)=\mu_j$ is the j-th moment of ρ. It is null if j is odd, since ρ is symmetric. Moreover, for any $j\in\mathbb{N}$,

$$\frac{\partial f_j}{\partial u}=f_{j+1}-f_jf_1\qquad\text{and}\qquad\frac{\partial f_j}{\partial v}=f_{j+2}-f_jf_2.$$

Therefore, for all $(x,y)\in A_I$,

$$\mathrm{D}^2_{(x,y)}I=\left(\mathrm{D}^2_{(u(x,y),v(x,y))}\Lambda\right)^{-1}=\begin{pmatrix}f_2-f_1^2 & f_3-f_1f_2\\ f_3-f_1f_2 & f_4-f_2^2\end{pmatrix}^{-1}(u(x,y),v(x,y)).$$

Denoting by $g=(f_2-f_1^2)(f_4-f_2^2)-(f_3-f_1f_2)^2$, the determinant of the positive definite symmetric matrix $\mathrm{D}^2\Lambda$, we get that for any $(x,y)\in A_I$,

$$\mathrm{D}^2_{(x,y)}I=\frac{1}{g(u(x,y),v(x,y))}\begin{pmatrix}f_4-f_2^2 & f_1f_2-f_3\\ f_1f_2-f_3 & f_2-f_1^2\end{pmatrix}(u(x,y),v(x,y)).$$

Moreover $(u(0,\sigma^2), v(0,\sigma^2)) = (0,0)$ thus

$$\frac{\partial u}{\partial x}(0,\sigma^2) = \frac{\partial^2 I}{\partial x^2}(0,\sigma^2) = \frac{f_4 - f_2^2}{g}(0,0) = \frac{\mu_4 - \sigma^4}{\sigma^2(\mu_4 - \sigma^4)} = \frac{1}{\sigma^2},$$

$$\frac{\partial v}{\partial y}(0,\sigma^2) = \frac{\partial^2 I}{\partial y^2}(0,\sigma^2) = \frac{f_2 - f_1^2}{g}(0,0) = \frac{\sigma^2}{\sigma^2(\mu_4 - \sigma^4)} = \frac{1}{\mu_4 - \sigma^4},$$

$$\frac{\partial u}{\partial y}(0,\sigma^2) = \frac{\partial v}{\partial x}(0,\sigma^2) = \frac{\partial^2 I}{\partial x \partial y}(0,\sigma^2) = \frac{f_1 f_2 - f_3}{g}(0,0) = 0.$$

Differentiating with respect to y, we get

$$\frac{\partial^3 I}{\partial y \partial x^2} = \frac{\partial u}{\partial y} \times \frac{\partial}{\partial u}\left(\frac{f_4 - f_2^2}{g}\right)(u,v) + \frac{\partial v}{\partial y} \times \frac{\partial}{\partial v}\left(\frac{f_4 - f_2^2}{g}\right)(u,v).$$

The first term of the addition, taken at $(0,\sigma^2)$, is null. For the second term, we need to compute the partial derivative of $(f_4 - f_2^2)/g$ with respect to v:

$$\begin{aligned}\frac{\partial}{\partial v}\left(\frac{f_4 - f_2^2}{g}\right) &= \frac{1}{g} \times \frac{\partial}{\partial v}\left(f_4 - f_2^2\right) - \frac{f_4 - f_2^2}{g^2} \times \frac{\partial g}{\partial v} \\ &= \frac{f_6 - 3f_2 f_4 + 2f_2^3}{g} - \frac{f_4 - f_2^2}{g^2} \times \frac{\partial g}{\partial v}.\end{aligned}$$

Let us differentiate with respect to v:

$$\begin{aligned}\frac{\partial g}{\partial v} &= f_2(f_6 - f_4 f_2) + f_4(f_4 - f_2^2) - f_1^2(f_6 - f_4 f_2) - 2f_4 f_1(f_3 - f_1 f_2) - 3f_2^2(f_4 - f_2^2) \\ &- 2f_3(f_5 - f_3 f_2) + 2f_1 f_2(f_5 - f_3 f_2) + 2f_2 f_3(f_3 - f_1 f_2) + 2f_1 f_3(f_4 - f_2^2).\end{aligned}$$

Taken at $(0,0)$, each term with even subscript vanishes and we have

$$\begin{aligned}\frac{\partial g}{\partial v}(0,0) &= \sigma^2(\mu_6 - \mu_4\sigma^2) + \mu_4(\mu_4 - \sigma^4) - 3\sigma^4(\mu_4 - \sigma^4) \\ &= \sigma^2\mu_6 - 3\mu_4\sigma^4 + 2\sigma^8 + (\mu_4 - \sigma^4)^2.\end{aligned}$$

Finally

$$\begin{aligned}\frac{\partial}{\partial v}\left(\frac{f_4 - f_2^2}{g}\right)(0,0) &= \frac{\mu_6 - 3\sigma^2\mu_4 + 2\sigma^6}{\sigma^2(\mu_4 - \sigma^4)} - \frac{\sigma^2\mu_6 - 3\mu_4\sigma^4 + 2\sigma^8 + (\mu_4 - \sigma^4)^2}{\sigma^4(\mu_4 - \sigma^4)} \\ &= \frac{\sigma^4 - \mu_4}{\sigma^4}.\end{aligned}$$

Therefore

$$\frac{\partial^3 I}{\partial y \partial x^2}(0,\sigma^2) = 0 + \frac{\partial v}{\partial y}(0,\sigma^2)\,\frac{\partial}{\partial v}\left(\frac{f_4 - f_2^2}{g}\right)(0,0) = \frac{1}{\mu_4 - \sigma^4} \times \frac{\sigma^4 - \mu_4}{\sigma^4} = -\frac{1}{\sigma^4}.$$

This is what we wanted to prove. Let us compute now the fourth partial derivative of I with respect to x. We have to obtain first an expression of the third partial derivative of I with respect to x:

$$\frac{\partial^3 I}{\partial x^3} = \frac{\partial u}{\partial x} \times \frac{\partial}{\partial u}\left(\frac{f_4 - f_2^2}{g}\right)(u,v) + \frac{\partial v}{\partial x} \times \frac{\partial}{\partial v}\left(\frac{f_4 - f_2^2}{g}\right)(u,v).$$

The only term we do not know is the partial derivative with respect to u of $(f_4 - f_2^2)/g$. We have

$$\begin{aligned}\frac{\partial}{\partial u}\left(\frac{f_4 - f_2^2}{g}\right) &= \frac{1}{g} \times \frac{\partial}{\partial u}\left(f_4 - f_2^2\right) - \frac{f_4 - f_2^2}{g^2} \times \frac{\partial g}{\partial u} \\ &= \frac{f_5 - f_4 f_1 - 2 f_2 f_3 + 2 f_2^2 f_1}{g} - \frac{f_4 - f_2^2}{g^2} \times \frac{\partial g}{\partial u},\end{aligned}$$

with

$$\begin{aligned}\frac{\partial g}{\partial u} = f_2(f_5 - f_4 f_1) + f_4(f_3 - f_2 f_1) - f_1^2(f_5 - f_4 f_1) - 2 f_4 f_1 (f_2 - f_1^2) \\ - 3 f_2^2 (f_3 - f_2 f_1) - 2 f_3 (f_4 - f_3 f_1) + 2 f_1 f_2 (f_4 - f_3 f_1) \\ + 2 f_2 f_3 (f_2 - f_1^2) + 2 f_1 f_3 (f_3 - f_2 f_1).\end{aligned}$$

Notice that this quantity vanishes at $(0,0)$. Therefore the partial derivative of $(f_4 - f_2^2)/g$ with respect to u, taken at $(0,0)$, is null as well and we get back that the third partial derivative of I with respect to x, taken at $(0,\sigma^2)$, is null. Differentiating once more, we obtain

$$\begin{aligned}\frac{\partial^4 I}{\partial x^4} = \frac{\partial u}{\partial x} \times \left(\frac{\partial u}{\partial x} \times \frac{\partial^2}{\partial u^2}\left(\frac{f_4 - f_2^2}{g}\right)(u,v) + \frac{\partial v}{\partial x} \times \frac{\partial^2}{\partial v \partial u}\left(\frac{f_4 - f_2^2}{g}\right)(u,v)\right) \\ + \frac{\partial^2 u}{\partial x^2} \times \frac{\partial}{\partial u}\left(\frac{f_4 - f_2^2}{g}\right)(u,v) + \frac{\partial^2 v}{\partial x^2} \times \frac{\partial}{\partial v}\left(\frac{f_4 - f_2^2}{g}\right)(u,v) \\ + \frac{\partial v}{\partial x} \times \left(\frac{\partial u}{\partial x} \times \frac{\partial^2}{\partial u \partial v}\left(\frac{f_4 - f_2^2}{g}\right)(u,v) + \frac{\partial v}{\partial x} \times \frac{\partial^2}{\partial v^2}\left(\frac{f_4 - f_2^2}{g}\right)(u,v)\right).\end{aligned}$$

Let us compute it at $(0,\sigma^2)$:

$$\frac{\partial^4 I}{\partial x^4}(0,\sigma^2) = \frac{1}{\sigma^2}\left(\frac{1}{\sigma^2}\frac{\partial^2}{\partial u^2}\left(\frac{f_4 - f_2^2}{g}\right)(0,0) + 0\right) + 0 + \frac{\sigma^4 - \mu_4}{\sigma^4}\frac{\partial^2 v}{\partial x^2}(0,\sigma^2) + 0,$$

with

$$\frac{\partial^2 v}{\partial x^2}(0,\sigma^2) = \frac{\partial}{\partial x}\left(\frac{\partial^2 I}{\partial x \partial y}\right)(0,\sigma^2) = \frac{\partial^3 I}{\partial x^2 \partial y}(0,\sigma^2) = -\frac{1}{\sigma^4}$$

and

$$\begin{aligned}\frac{\partial^2}{\partial u^2}\left(\frac{f_4 - f_2^2}{g}\right) = \frac{1}{g}\frac{\partial^2}{\partial u^2}(f_4 - f_2^2) - \frac{2}{g^2}\frac{\partial g}{\partial u}\frac{\partial}{\partial u}(f_4 - f_2^2) - \frac{f_4 - f_2^2}{g^2}\frac{\partial^2 g}{\partial u^2} \\ + \frac{2}{g^3}\left(\frac{\partial g}{\partial u}\right)^2 (f_4 - f_2^2).\end{aligned}$$

Hence

$$\frac{\partial^2}{\partial u^2}\left(\frac{f_4 - f_2^2}{g}\right)(0,0) = \frac{1}{\sigma^4(\mu_4 - \sigma^4)}\left(\sigma^2 \frac{\partial^2}{\partial u^2}(f_4 - f_2^2)(0,0) - \frac{\partial^2 g}{\partial u^2}(0,0)\right).$$

The two remaining terms are the derivatives of quantities which we have already computed. We evaluate them directly at $(0,0)$, which is straightforward since $f_j(0,0) = 0$ when j is odd:

$$\frac{\partial^2}{\partial u^2}(f_4 - f_2^2)(0,0) = \frac{\partial}{\partial u}(f_5 - f_4 f_1 - 2 f_2 f_3 + 2 f_2^2 f_1)(0,0) = \mu_6 - 3\sigma^2 \mu_4 + 2\sigma^6$$

and

$$\frac{\partial^2 g}{\partial u^2}(0,0) = \frac{\partial}{\partial u}\left(\frac{\partial g}{\partial u}\right)(0,0) = \sigma^2(\mu_6 - \mu_4\sigma^2) + \mu_4(\mu_4 - \sigma^4) - 0 - 2\mu_4\sigma^4 \\ - 3\sigma^4(\mu_4 - \sigma^4) - 2\mu_4^2 + 2\sigma^4\mu_4 + 2\sigma^4\mu_4 + 0.$$

This is equal to $\sigma^2\mu_6 - \mu_4^2 + 3\sigma^8 - 3\mu_4\sigma^4$ after simplification. Thus we have

$$\frac{\partial^2}{\partial u^2}\left(\frac{f_4 - f_2^2}{g}\right)(0,0) = \frac{\sigma^2\mu_6 - 3\sigma^4\mu_4 + 2\sigma^8 - \sigma^2\mu_6 + \mu_4^2 - 3\sigma^8 + 3\mu_4\sigma^4}{\sigma^4(\mu_4 - \sigma^4)} \\ = \frac{\mu_4^2 - \sigma^8}{\sigma^4(\mu_4 - \sigma^4)} = \frac{\mu_4 + \sigma^4}{\sigma^4}.$$

Finally

$$\frac{\partial^2 I}{\partial x^4}(0,\sigma^2) = \frac{\mu_4 + \sigma^4}{\sigma^8} - \frac{\sigma^4 - \mu_4}{\sigma^8} = \frac{2\mu_4}{\sigma^8}.$$

We obtain the announced term and the proof is completed. □

7 Proof of theorem 2

We first give conditions on the probability measure ρ in order to apply theorem A.5 (see appendix A) to the distribution ν_ρ. We will use then Laplace's method, as we announced in the heuristics of section 3, to obtain the fluctuations theorem 2. The proof will rely on the expansion of $I - F$ around $(0,\sigma^2)$ given in proposition 15. We will also use the variant of Varadhan's lemma, stated in proposition 14. We start with the following lemma:

Lemma 16. *If ρ has a probability density f with respect to the Lebesgue measure on $\mathbb{R}$, then ν_ρ^{*2} has the density*

$$f_2 : (x,y) \longmapsto \frac{1}{\sqrt{2y - x^2}}\, f\left(\frac{x + \sqrt{2y - x^2}}{2}\right) f\left(\frac{x - \sqrt{2y - x^2}}{2}\right) \mathbb{1}_{x^2 < 2y}$$

with respect to the Lebesgue measure on $\mathbb{R}^2$.

Proof. Let h be a bounded continuous function from $\mathbb{R}^2$ to $\mathbb{R}$. We have

$$\int_{\mathbb{R}^2} h(x,y)\, d\nu_\rho^{*2}(x,y) = \int_{\mathbb{R}^2} h((z,z^2) + (t,t^2))\, d\rho(z)\, d\rho(t) \\ = \int_{D^+} h(z+t, z^2+t^2) f(z) f(t)\, dz\, dt + \int_{D^-} h(z+t, z^2+t^2) f(z) f(t)\, dz\, dt,$$

with $D^+ = \{\, (z,t) \in \mathbb{R}^2 : z > t \,\}$ and $D^- = \{\, (z,t) \in \mathbb{R}^2 : z < t \,\}$. Indeed, the Lebesgue measure of the set $\{\, (z,t) \in \mathbb{R}^2 : z = t \,\}$ is null. Let us denote respectively by I_+ and I_- the two previous integrals.

We define $\phi : (z,t) \in \mathbb{R}^2 \longmapsto (u,v) = (z+t, z^2+t^2)$. It is a one to one map from D^+ (resp. from D^-) onto $\Delta_2 = \{\, (u,v) \in \mathbb{R}^2 : u^2 < 2v \,\}$. Moreover ϕ is $\mathcal{C}^1$ on $D^+ \cup D^-$ and its Jacobian in (z,t) is $2|z - t| = 2\sqrt{2v - u^2} \neq 0$. The change of variables given by ϕ yields

$$I_+ = I_- = \int_{\Delta_2} h(u,v) \frac{1}{2\sqrt{2v - u^2}}\, f\left(\frac{u + \sqrt{2v - u^2}}{2}\right) f\left(\frac{u - \sqrt{2v - u^2}}{2}\right) du\, dv.$$

By adding theses two terms, we get the lemma. $\square$

By theorem A.5 in appendix, the expansion of g_n holds as soon as there exists $q \in [1,+\infty[$ such that $\widehat{f_2} \in \mathrm{L}^q(\mathbb{R}^d)$. However the computation of $\widehat{f_2}$ is not feasible in general. Proposition A.6 says that the previous condition is satisfied if there exists $p \in]1,2]$ such that $f_2 \in \mathrm{L}^p(\mathbb{R}^d)$ so that the expansion is true. Let us take a look at this:

$$\int_{\mathbb{R}^2} |f_2(u,v)|^p \, du\, dv$$
$$= \int_{\mathbb{R}^2} \frac{f^p\Big((u+\sqrt{2v-u^2})/2\Big)\, f^p\Big((u-\sqrt{2v-u^2})/2\Big)}{(2v-u^2)^{p/2}} \mathbb{1}_{u^2<2v}\, du\, dv.$$

Let us make the change of variables given by

$$(u,v) \longmapsto (x,y) = \frac{1}{2}(u+\sqrt{2v-u^2}, u+\sqrt{2v-u^2}),$$

which is a $\mathcal{C}^1$-diffeomorphism from Δ_2 to D^+ (see the proof of the previous lemma) with Jacobian in (u,v), $2\sqrt{2v-u^2} = 2(y-x) > 0$:

$$\int_{\mathbb{R}^2} |f_2(u,v)|^p \, du\, dv = \int_{\mathbb{R}^2} \frac{f^p(x)f^p(y)}{(y-x)^p} 2(y-x)\mathbb{1}_{y>x}\, dx\, dy.$$

By symmetry in x and y, we get

$$\int_{\mathbb{R}^2} |f_2(u,v)|^p \, du\, dv = \int_{\mathbb{R}^2} f^p(x)f^p(y)|y-x|^{1-p}\, dx\, dy.$$

Then we get the following proposition :

Proposition 17. *Suppose that ρ has a density f with respect to the Lebesgue measure on $\mathbb{R}$ such that, for some $p \in]1,2]$,*

$$\int_{\mathbb{R}^2} f^p(x+y)f^p(y)|x|^{1-p}\, dx\, dy < +\infty.$$

Then, for n large enough, $\widetilde{\nu}_{n,\rho}$ has a density g_n with respect to the Lebesgue measure on $\mathbb{R}^2$ such that, for any compact subset K_I of A_I, when $n \to +\infty$, uniformly over $(x,y) \in K_I$.

$$g_n(x,y) \sim \frac{n}{2\pi}\left(\det \mathrm{D}^2_{(x,y)} I\right)^{1/2} e^{-nI(x,y)}.$$

Let us prove now theorem 2. Suppose that ρ is a probability measure on $\mathbb{R}$ with an even density f such that there exist $v_0 > 0$ and $p \in]1,2]$ such that

$$\int_{\mathbb{R}} e^{v_0 z^2} f(z)\, dz < +\infty \qquad \text{and} \qquad \int_{\mathbb{R}^2} f^p(x+y)f^p(y)|x|^{1-p}\, dx\, dy < +\infty.$$

The first inequality implies that $\mathbb{R}\times]-\infty, v_0[\subset D_\Lambda$ and thus $(0,0) \in \overset{o}{D}_\Lambda$. Moreover ρ is symmetric (since f is even) and its support contains at least three points (since ρ has a density). Proposition 15 implies that there exists $\delta > 0$ such that

$$\forall (x,y) \in \mathrm{B}_\delta \qquad G(x,y) = I(x,y) - \frac{x^2}{2y} \geq \frac{(y-\sigma^2)^2}{4(\mu_4-\sigma^4)} + \frac{\mu_4 x^4}{24\sigma^8}, \qquad (*)$$

where μ_4 denotes the fourth moment of ρ and B_δ the open ball of radius δ centered at $(0,\sigma^2)$. We can reduce δ, in order to have $\mathrm{B}_\delta \subset K_I$ where K_I is a compact subset of A_I. Moreover $A_I \subset \overset{o}{D}_I \subset \Delta^*$ thus $\mathrm{B}_\delta \cap \Delta^* = \mathrm{B}_\delta$.

Let $n \in \mathbb{N}$ and let $f : \mathbb{R} \longrightarrow \mathbb{R}$ be a bounded continuous function. We have

$$\mathbb{E}_{\tilde{\mu}_{n,\rho}}\left(f\left(\frac{S_n}{n^{3/4}}\right)\right) = \frac{1}{Z_n}\int_{\Delta^*} f(xn^{1/4})\exp\left(\frac{nx^2}{2y}\right)\, d\widetilde{\nu}_{n,\rho}(x,y) = \frac{A_n+B_n}{Z_n},$$

with

$$A_n = \int_{\mathrm{B}_\delta} f(xn^{1/4})\exp\left(\frac{nx^2}{2y}\right)\, d\widetilde{\nu}_{n,\rho}(x,y),$$

$$B_n = \int_{\Delta^*\cap \mathrm{B}_\delta^c} f(xn^{1/4})\exp\left(\frac{nx^2}{2y}\right)\, d\widetilde{\nu}_{n,\rho}(x,y).$$

Let us introduce $e^{-nI(x,y)}$ in the expression of A_n, in order to use proposition 17 :

$$A_n = n\int_{\mathrm{B}_\delta} f(xn^{1/4})e^{-nG(x,y)}H_n(x,y)\,dx\,dy,$$

where we set $H_n(x,y) = e^{nI(x,y)}g_n(x,y)/n$. We define

$$\mathrm{B}_{\delta,n} = \{\,(x,y)\in\mathbb{R}^2 : x^2/\sqrt{n} + y^2/n \le \delta^2\,\}.$$

Let us make the change of variables given by $(x,y) \longmapsto (xn^{-1/4}, yn^{-1/2}+\sigma^2)$, with Jacobian $n^{-3/4}$:

$$A_n = n^{1/4}\int_{\mathrm{B}_{\delta,n}} f(x)\exp\left(-nG\left(\frac{x}{n^{1/4}}, \frac{y}{\sqrt{n}}+\sigma^2\right)\right) H_n\left(\frac{x}{n^{1/4}}, \frac{y}{\sqrt{n}}+\sigma^2\right)\,dx\,dy.$$

We check now that we can apply the dominated convergence theorem to this integral. The uniform expansion of g_n (see proposition 17) means that for any $\alpha > 0$, there exists $n_0 \in \mathbb{N}$ such that

$$(x,y)\in K_I \quad\text{and}\quad n \ge n_0 \quad\Longrightarrow\quad \left| H_n(x,y)\,2\pi\left(\det \mathrm{D}^2_{(x,y)}I\right)^{-1/2} - 1\right| \le \alpha.$$

If $(x,y)\in\mathrm{B}_{\delta,n}$, then $(x_n,y_n) = (xn^{-1/4}, yn^{-1/2}+\sigma^2)\in\mathrm{B}_\delta\subset K_I$, thus for all $n\ge n_0$ and $(x,y)\in\mathrm{B}_{\delta,n}$,

$$\left| H_n\left(\frac{x}{n^{1/4}}, \frac{y}{\sqrt{n}}+\sigma^2\right) 2\pi\left(\det \mathrm{D}^2_{(x_n,y_n)}I\right)^{-1/2} - 1\right| \le \alpha.$$

Moreover $(x_n,y_n)\to(0,\sigma^2)$ thus, by continuity,

$$\left(\mathrm{D}^2_{(x_n,y_n)}I\right)^{-1/2} \underset{n\to+\infty}{\longrightarrow} \left(\mathrm{D}^2_{(0,\sigma^2)}I\right)^{-1/2} = \left(\mathrm{D}^2_{(0,0)}\Lambda\right)^{1/2},$$

whose determinant is equal to $\sqrt{\sigma^2(\mu_4-\sigma^4)}$. Therefore

$$\mathbb{1}_{\mathrm{B}_{\delta,n}}(x,y)H_n\left(\frac{x}{n^{1/4}}, \frac{y}{\sqrt{n}}+\sigma^2\right) \underset{n\to+\infty}{\longrightarrow} \left(4\pi^2\sigma^2(\mu_4-\sigma^4)\right)^{-1/2}.$$

The expansion of G in the neighbourhood of $(0,\sigma^2)$ implies that

$$\exp\left(-nG\left(\frac{x}{n^{1/4}},\frac{y}{\sqrt{n}}+\sigma^2\right)\right)\underset{n\to+\infty}{\longrightarrow}\exp\left(-\frac{y^2}{2(\mu_4-\sigma^4)}-\frac{\mu_4x^4}{12\sigma^8}\right).$$

Let us check that the integrand is dominated by an integrable function, which is independent of n. The function

$$(x,y)\longmapsto\left(\mathrm{D}^2_{(x,y)}I\right)^{-1/2}$$

is bounded on B_δ by some $M_\delta>0$. The uniform expansion of g_n implies that for all $(x,y)\in\mathrm{B}_\delta$, $H_n(x,y)\leq C_\delta$ for some constant $C_\delta>0$. Finally, the inequality $(*)$ above yields that

$$\mathbb{1}_{\mathrm{B}_{\delta,n}}(x,y)f(x)\exp\left(-nG\left(\frac{x}{n^{1/4}},\frac{y}{\sqrt{n}}+\sigma^2\right)\right)H_n\left(\frac{x}{n^{1/4}},\frac{y}{\sqrt{n}}+\sigma^2\right)$$
$$\leq\|f\|_\infty C_\delta\exp\left(-\frac{y^2}{4(\mu_4-\sigma^4)}-\frac{\mu_4x^4}{24\sigma^8}\right).$$

The right term is an integrable function on $\mathbb{R}^2$, thus it follows from the dominated convergence theorem that

$$A_n\underset{+\infty}{\sim}n^{1/4}\int_{\mathbb{R}^2}f(x)\frac{1}{\sqrt{2\pi\sigma^2}\sqrt{2\pi(\mu_4-\sigma^4)}}\exp\left(-\frac{y^2}{2(\mu_4-\sigma^4)}-\frac{\mu_4x^4}{12\sigma^8}\right)dx\,dy.$$

By Fubini's theorem, we get

$$A_n\underset{+\infty}{\sim}\frac{n^{1/4}}{\sqrt{2\pi\sigma^2}}\int_{\mathbb{R}}f(x)\exp\left(-\frac{\mu_4x^4}{12\sigma^8}\right)\,dx.$$

Let us focus now on B_n. The distribution ρ is symmetric, it has a density and $(0,0)$ belongs to the interior of D_Λ, thus proposition 14 implies that there exist $\varepsilon>0$ and $n_0\geq1$ such that for any $n\geq n_0$,

$$\int_{\Delta^*\cap\mathrm{B}_\delta^c}\exp\left(\frac{nx^2}{2y}\right)\,d\widetilde{\nu}_{n,\rho}(x,y)\leq e^{-n\varepsilon},$$

and thus $B_n\leq\|f\|_\infty e^{-n\varepsilon}$, so that $B_n=o(n^{1/4})$. Therefore

$$A_n+B_n\underset{+\infty}{\sim}\frac{n^{1/4}}{\sqrt{2\pi\sigma^2}}\int_{\mathbb{R}}f(x)\exp\left(-\frac{\mu_4x^4}{12\sigma^8}\right)\,dx.$$

Applying this to $f=1$, we get

$$Z_n\underset{+\infty}{\sim}\frac{2n^{1/4}}{\sqrt{2\pi\sigma^2}}\int_0^{+\infty}\exp\left(-\frac{\mu_4x^4}{12\sigma^8}\right)\,dx=\frac{n^{1/4}}{\sqrt{2\pi\sigma^2}}\frac{1}{2}\left(\frac{12\sigma^8}{\mu_4}\right)^{1/4}\Gamma\left(\frac{1}{4}\right),$$

where we made the change of variables $y=\mu_4x^4/(12\sigma^8)$. Finally

$$\mathbb{E}_{\tilde{\mu}_{n,\rho}}\left(f\left(\frac{S_n}{n^{3/4}}\right)\right)\underset{+\infty}{\sim}\left(\frac{4\mu_4}{3\sigma^8}\right)^{1/4}\Gamma\left(\frac{1}{4}\right)^{-1}\int_{\mathbb{R}}f(x)\exp\left(-\frac{\mu_4x^4}{12\sigma^8}\right)\,dx.$$

The ultimate change of variables $s=\mu_4^{1/4}x/\sigma^2$ gives us theorem 2.

Appendix A
General results on the Cramér transform

This appendix presents some general results on the Cramér transform of a probability distribution on $\mathbb{R}^d$.

A probability measure $\mathbb{R}$ is said to be degenerate if it is a Dirac mass. The following definition generalizes this notion for measures on $\mathbb{R}^d$:

Definition A.1. *A probability measure ν on $\mathbb{R}^d$, $d \geq 2$, is said to be degenerate if its support is included in a hyperplane of $\mathbb{R}^d$, i.e., there exists a hyperplane $\mathcal{H}$ of $\mathbb{R}^d$ such that $\nu(\mathcal{H}) = 1$.*

A first consequence of the non-degeneracy of ν is that its covariance matrix is a symmetric positive definite matrix (see section III.5 of [15] for a proof).

From now onwards, we consider ν a non-degenerate probability measure on $\mathbb{R}^d$. The Log-Laplace L of ν is defined in $\mathbb{R}^d$ by

$$\forall \lambda \in \mathbb{R}^d \qquad L(\lambda) = \ln \int_{\mathbb{R}^d} e^{\langle \lambda, z \rangle} \, d\nu(z),$$

where $\langle \ , \ \rangle$ denotes the inner product in $\mathbb{R}^d$. It is a convex function on $\mathbb{R}^d$ which takes its values in $]-\infty, +\infty]$. The Fenchel-Legendre transform of L is called the Cramér transform of ν and is defined on $\mathbb{R}^d$ by

$$\forall x \in \mathbb{R}^d \qquad J(x) = \sup_{\lambda \in \mathbb{R}^d} \left(\langle \lambda, x \rangle - L(\lambda) \right).$$

It is a non-negative, convex and lower semi-continuous function. We denote by D_L and D_J the convex sets where L and J are finite. Notice that, if $\overset{o}{D}_L \neq \varnothing$, then L is $\mathcal{C}^\infty$ on $\overset{o}{D}_L$. We refer to section 2.2 of [10], section VII.5 of [13] and sections 25 and 26 of [21] for the main results on L and J Cramér's theorem (theorem B.4 in appendix) links J and the large deviations of S_n/n where S_n is the sum of n independent random variables with common distribution ν.

We are interested in the points λ realizing the supremum defining $J(x)$, for $x \in D_J$. We denote by $\mathcal{C}$ the closed convex hull of the support of ν.

Lemma A.2. *Let ν be a non-degenerate probability measure on $\mathbb{R}^d$. The interior of $\mathcal{C}$ is not empty and $\overset{o}{\mathcal{C}} \subset D_J \subset \mathcal{C}$. Moreover for any $x \in \overset{o}{\mathcal{C}}$, the supremum defining $J(x)$ is realized for some value $\lambda(x) \in D_L$.*

Proof. The non-degeneracy of ν means that its support is not included in a hyperplane of $\mathbb{R}^d$. Therefore the support of ν contains d linearly independent vectors and the interior of the convex hull of these vectors is non-empty. Thus $\overset{o}{\mathcal{C}}$ is non-empty.

Suppose that $\mathcal{C} \neq \mathbb{R}^d$ (otherwise it is immediate that $D_J \subset \mathcal{C}$). Let $x \notin \mathcal{C}$. By the Hahn-Banach theorem, there exists $\lambda \in \mathbb{R}^d$ and $a \in \mathbb{R}$ such that

$$\forall y \in \mathcal{C} \qquad \langle \lambda, y \rangle \leq a < \langle \lambda, x \rangle.$$

Since $\nu(\mathcal{C}) = 1$, Jensen's inequality implies that

$$\forall t > 0 \qquad J(x) \geq -\ln \int_{\mathcal{C}} \exp(t\langle \lambda, y \rangle - t\langle \lambda, x \rangle) \, d\nu(y) \geq t(\langle \lambda, x \rangle - a).$$

Sending t to $+\infty$, we conclude that $J(x)=+\infty$. Thus $D_J \subset \mathcal{C}$.

Let $x \in \overset{o}{\mathcal{C}}$ and let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $\mathbb{R}^d$ such that

$$\begin{aligned}J(x) &= \lim_{n\to+\infty}\Big(\langle\lambda_n, x\rangle - \ln\int_{\mathbb{R}^d}\exp(\langle\lambda_n, z\rangle)\,d\nu(z)\Big)\\ &= -\ln\lim_{n\to+\infty}\int_{\mathbb{R}^d}\exp(\langle\lambda_n, z-x\rangle)\,d\nu(z).\end{aligned}$$

We suppose that $|\lambda_n| \to +\infty$ and we show that it leads to a contradiction. For all $n\in\mathbb{N}$, we set $u_n = \lambda_n|\lambda_n|^{-1}$. Then $(u_n)_{n\in\mathbb{N}}$ is a bounded sequence. Thus, up to the extraction of a subsequence, we might assume that it converges to some vector $u\in\mathbb{R}^d$ whose norm is 1. Let v belong to the support of ν and let U be an open subset of $\mathbb{R}^d$ containing v. We have then $\nu(U)>0$. Suppose that for any $z\in U$, $\langle u, z-x\rangle > 0$. Then, by Fatou's lemma,

$$\begin{aligned}+\infty &= \int_U \liminf_{n\to+\infty}\exp(|\lambda_n|\langle u_n, z-x\rangle)\,d\nu(z)\\ &\leq \liminf_{n\to+\infty}\int_U \exp(|\lambda_n|\langle u_n, z-x\rangle)\,d\nu(z).\end{aligned}$$

Hence

$$\exp(-J(x)) = \lim_{n\to+\infty}\int_{\mathbb{R}^d}\exp(|\lambda_n|\langle u_n, z-x\rangle)\,d\nu(z) = +\infty.$$

Thus $J(x)=-\infty$, which is absurd since J is a non-negative function. We conclude that for all v in the support of ν and for any open subset U of $\mathbb{R}^d$ containing v, there exists $z\in U$ such that $\langle u, z-x\rangle \leq 0$. It follows that, for any v in the support of ν, $\langle u, v\rangle \leq \langle u, x\rangle$. This inequality is stable by convex combinations, thus

$$\forall y\in\mathcal{C}\qquad \langle u, y\rangle \leq \langle u, x\rangle.$$

Since $x\in\overset{o}{\mathcal{C}}$, there exists a ball B_x centered at x and contained in $\mathcal{C}$. Thus there exists $y_0\in\mathrm{B}_x$ such that $\langle u, y_0\rangle > \langle u, x\rangle$, which is absurd. Therefore $(\lambda_n)_{n\in\mathbb{N}}$ is a bounded sequence. Hence there exists a subsequence $(\lambda_{\phi(n)})_{n\in\mathbb{N}}$ and $\lambda(x)\in\mathbb{R}^d$ such that $\lambda_{\phi(n)}\to\lambda(x)$. By Fatou's lemma,

$$\begin{aligned}J(x) &= \langle\lambda(x), x\rangle - \ln\lim_{n\to+\infty}\int_{\mathbb{R}^d}\exp(\langle\lambda_n, z\rangle)\,d\nu(z)\\ &\leq \langle\lambda(x), x\rangle - \ln\int_{\mathbb{R}^d}\liminf_{n\to+\infty}\exp(\langle\lambda_n, z\rangle)\,d\nu(z)\\ &= \langle\lambda(x), x\rangle - \ln\int_{\mathbb{R}^d}\exp(\langle\lambda(x), z\rangle)\,d\nu(z) \leq J(x).\end{aligned}$$

Thus $J(x) = \langle\lambda(x), x\rangle - L(\lambda(x))$. Since $L(\lambda(x)) \neq -\infty$, this formula implies that $J(x) < +\infty$ and thus that $\overset{o}{\mathcal{C}} \subset D_J$. Moreover if $L(\lambda(x)) = +\infty$ then $J(x) = -\infty$, which is absurd. Therefore $L(\lambda(x)) < \infty$. This shows that the supremum defining $J(x)$ is realized at a point $\lambda(x)$ such that $\Lambda(\lambda(x)) < +\infty$. $\square$

If D_L is an open subset of $\mathbb{R}^d$ then for all $(x,y)\in\overset{o}{D}_J = \overset{o}{\mathcal{C}}$, the supremum defining $J(x)$ is realized at some $\lambda(x)\in\overset{o}{D}_L$. This is the case when the support of ν is bounded, and also for the distribution ν_ρ when ρ is the Gaussian $\mathcal{N}(0,\sigma^2)$, where we have then $D_L = \mathbb{R}\times\,]-\infty, 1/(2\sigma^2)[$.

Now we study the smoothness of J.

Notation. *If f is a differentiable function on an open subset U of $\mathbb{R}^d$, we denote by $\mathrm{D}_x f$ the differential of f at $x \in U$. If f is real-valued, we denote:*
$\star$ $\mathrm{D}^2_x f$ *its second differential at $x \in U$ (considered as a matrix of size $d \times d$).*
$\star$ ∇f *the function $U \longrightarrow \mathbb{R}^d$ such that*

$$\forall x \in U \quad \forall y \in \mathbb{R}^d \qquad \langle \nabla f(x), y \rangle = \mathrm{D}_x f(y).$$

We define the admissible domain of J:

Definition A.3. *Let ν be a non-degenerate probability measure on $\mathbb{R}^d$ such that the interior of D_L is non-empty. The admissible domain of J is the set $A_J = \nabla L\,(\overset{o}{D}_L)$.*

The following proposition states that A_J, the admissible domain of J, is an open subset of $\mathbb{R}^d$, and that J is $\mathcal{C}^\infty$ on A_J.

Proposition A.4. *Let ν be a non-degenerate probability measure on $\mathbb{R}^d$ such that the interior of D_L is non-empty. Let A_J be the admissible domain of J.*
(a) *The function ∇L is a $\mathcal{C}^\infty$-diffeomorphism from $\overset{o}{D}_L$ to A_J. Moreover*

$$A_J \subset D_J = \{\, x \in \mathbb{R}^d : J(x) < +\infty \,\}.$$

(b) *Denote by λ the inverse $\mathcal{C}^\infty$-diffeomorphism of ∇L. Then the map J is $\mathcal{C}^\infty$ on A_J and for any $x \in A_J$,*

$$J(x) = \langle x, \lambda(x) \rangle - L(\lambda(x)),$$

$$\nabla J(x) = (\nabla L)^{-1}(x) = \lambda(x) \qquad \text{and} \qquad \mathrm{D}^2_x J = \left(\mathrm{D}^2_{\lambda(x)} L \right)^{-1}.$$

(c) *If D_L is an open subset of $\mathbb{R}^d$ then $A_J = \overset{o}{D}_J = \overset{o}{\mathcal{C}}$ where $\mathcal{C}$ denotes the convex hull of the support of ν.*

Proof. The points (a) and (b) are proved in section 2 of [1], section 1.5 of [7] and section 26 of [21] (see also section VIII.4 of [13] in the case $D_L = \mathbb{R}^d$). Let us prove the point (c). If D_L is an open subset of $\mathbb{R}^d$ then lemma A.2 implies that for $x \in \overset{o}{\mathcal{C}} = \overset{o}{D}_J$, the supremum defining $J(x)$ is realized at some point $\lambda(x) \in D_L = \overset{o}{D}_L$. The function L is differentiable at $\lambda(x)$ and the point (b) yields that

$$x = \nabla L(\lambda(x)) \in \Lambda(\overset{o}{D}_L) = A_J.$$

Thus $\overset{o}{D}_J \subset A_J$. Finally, since $A_J \subset D_J$ and A_J is open, we have $A_J = \overset{o}{D}_J = \overset{o}{\mathcal{C}}$. This proves (c). □

Let ν be a probability distribution on $\mathbb{R}^d$ having a density with respect to the Lebesgue measure and let S_n be the sum of n independent and identically distributed random variables with distribution ν. The following theorem states that, under some hypothesis allowing the Fourier inversion, the density of the distribution of S_n/n is asymptotically a function of J, the Cramér transform of ν. We refer to section 3 of the article of C. Andriani and P. Baldi [1] for a proof.

Theorem A.5. *Let ν be a non-degenerate probability measure on $\mathbb{R}^d$. We denote by L its Log-Laplace and by J its Cramér transform. Suppose that $\overset{o}{D}_L \neq \varnothing$ and that there exists $n_0 \geq 1$ such that*

$$\widehat{\nu^{*n_0}} \in \mathrm{L}^1(\mathbb{R}^d).$$

We denote by A_J the admissible domain of J. Let $(X_n)_{n\geq 1}$ be a sequence of independent and identically distributed random variables with distribution ν. For any $n \geq n_0$, the random variable $\overline{X}_n = (X_1 + \cdots + X_n)/n$ has a density g_n with respect to the Lebesgue measure on $\mathbb{R}^d$. If K_J is a compact subset of A_J then, uniformly over $x \in K_J$, when n goes to $+\infty$,

$$g_n(x) \sim \left(\frac{n}{2\pi}\right)^{d/2} \left(\det \mathrm{D}_x^2 J\right)^{1/2} e^{-nJ(x)}.$$

Proposition A.6. *Let ν be a non-degenerate probability measure on $\mathbb{R}^d$ such that $\overset{o}{D}_L \neq \varnothing$. If there exists $m \in \mathbb{N}$ and $p \in]1, 2]$ such that ν^{*m} has a density $f_m \in \mathrm{L}^p(\mathbb{R}^d)$ then the hypothesis of theorem A.5 are verified.*

Proof. The Hausdorff-Young inequality (see theorem 1.2.1 of [6]) implies that $\widehat{f_m} \in \mathrm{L}^r(\mathbb{R}^d)$, with $r = p/(p-1)$. Moreover $\widehat{f_m}$ is bounded thus $\widehat{f_m} \in \mathrm{L}^q(\mathbb{R}^d)$, where q is a positive integer larger than r. Therefore

$$\widehat{\nu^{*mq}} = \left(\widehat{\nu^{*m}}\right)^q = \left(\widehat{f_m}\right)^q \in \mathrm{L}^1(\mathbb{R}^d).$$

Hence the hypothesis of the theorem are verified with $n_0 = mq$. □

Appendix B
Some results on large deviations

Let $(\mathcal{X}, \mathcal{B})$ be a topological space. We refer to the section 1.2 of [10] for the two following definitions :

Definition B.1. *A rate function on $\mathcal{X}$ is a non-negative map J defined on $\mathcal{X}$ and which is lower semi-continuous, that is, for any $\alpha > 0$, the level set*

$$\{\, x \in \mathcal{X} : J(x) \leq \alpha \,\}$$

is a closed subset of $\mathcal{X}$. A good rate function is a rate function for which all these level sets are compact sets of $\mathcal{X}$.

Definition B.2. *A sequence $(\mu_n)_{n\geq 1}$ of probability measures on $\mathcal{X}$ satisfies a large deviation principle with speed n and governed by the rate function J if, for any $A \in \mathcal{B}$,*

$$-\inf \{\, J(x) : x \in \overset{o}{A} \,\} \leq \liminf_{n\to+\infty} \frac{1}{n} \ln \mu_n(A)$$
$$\leq \limsup_{n\to+\infty} \frac{1}{n} \ln \mu_n(A) \leq -\inf \{\, J(x) : x \in \overline{A} \,\}.$$

The following lemma is a variant of the upper bound of Varadhan's lemma (see lemma 4.3.6 of [10]).

Lemma B.3. *Let $\mathcal{X}$ be a regular topological space endowed with its Borel σ-field $\mathcal{B}$. Let $(\nu_n)_{n\geq 1}$ be a sequence of probability measures defined on $(\mathcal{X}, \mathcal{B})$ which satisfies a large deviation principle with speed n, governed by the good* *rate function J. For any bounded continuous function $f : \mathcal{X} \longrightarrow \mathbb{R}$, we have for any closed subset A of $\mathcal{X}$,*

$$\limsup_{n\to+\infty} \frac{1}{n} \ln \int_A e^{nf(x)}\, d\nu_n(x) \leq \sup_{x\in A} \left(f(x) - J(x)\right).$$

We end this appendix with the Cramér theorem in $\mathbb{R}^d$ (see theorem 2.2.30 of [10]) :

Theorem B.4 (Cramér)**.** *Let ν be a probability measure on $\mathbb{R}^d$, $d \geq 1$. We denote by L its Log-Laplace and by J its Cramér transform. Let $(X_n)_{n\geq 1}$ be a sequence of independent random variables with common law ν. We define*

$$\forall n \geq 1 \qquad S_n = X_1 + \cdots + X_n.$$

If L is finite in the neighbourhood of 0 then J is a good rate function and the sequence of the laws of S_n/n, $n \geq 1$, satisfies the large deviation principle with speed n and governed by J.

References

[1] Cristina Andriani and Paolo Baldi. Sharp estimates of deviations of the sample mean in many dimensions. *Ann. Inst. H. Poincaré Probab. Statist.*, 33(3):371–385, 1997.

[2] Markus Josef Aschwanden (Editor). *Self-Organized Criticality Systems.* Open Academic Press, 2013.

[3] Per Bak. *How nature works.* Copernicus, 1996. The science of self-organized criticality.

[4] Per Bak, Chao Tang, and Kurt Wiesenfeld. Self-organized criticality: An explanation of 1/f noise. *Phys. Rev. Lett.*, 59:381–384, 1987.

[5] Per Bak, Chao Tang, and Kurt Wiesenfeld. Self-organized criticality. *Phys. Rev. A (3)*, 38(1):364–374, 1988.

[6] Jöran Bergh and Jörgen Löfström. *Interpolation spaces. An introduction.* Springer-Verlag, 1976. Grundlehren der Mathematischen Wissenschaften, No. 223.

[7] A. A. Borovkov and A. A. Mogulskii. Large deviations and testing statistical hypotheses. I. Large deviations of sums of random vectors. *Siberian Adv. Math.*, 2(3):52–120, 1992. Siberian Advances in Mathematics.

[8] Raphaël Cerf. *On Cramér's theory in infinite dimensions*, volume 23 of *Panoramas et Synthèses.* Société Mathématique de France, 2007.

[9] Jan De Boer, Bernard Derrida, Henrik Flyvbjerg, Andrew D. Jackson, and Tilo Wettig. A simple model of self-organized biological evolution. *Phys. Rev. Lett.*, (73):906–909, 1994.

[10] Amir Dembo and Ofer Zeitouni. *Large deviations techniques and applications*, volume 38 of *Stochastic Modelling and Applied Probability.* Springer-Verlag, 2010.

[11] Deepak Dhar. Theoretical studies of self-organized criticality. *Phys. A*, 369(1):29–70, 2006.

[12] Theodor Eisele and Richard S. Ellis. Multiple phase transitions in the generalized Curie-Weiss model. *J. Statist. Phys.*, 52(1-2):161–202, 1988.

[13] Richard S. Ellis. *Entropy, large deviations, and statistical mechanics.* Classics in Maths. Springer-Verlag, 2006.

[14] Richard S. Ellis and Charles M. Newman. Limit theorems for sums of dependent random variables occurring in statistical mechanics. *Z. Wahrsch. Verw. Gebiete*, 44(2):117–139, 1978.

[15] William Feller. *An introduction to probability theory and its applications. Vol. II.* Second edition. John Wiley & Sons Inc., 1971.

[16] Matthias Gorny. Le modèle d'Ising Curie-Weiss et ses généralisations. *Master Thesis, Université Paris Sud*, 2012.

[17] Matthias Gorny. A Curie-Weiss model of Self-Organized Criticality : The Gaussian Case. *Markov processes and related fields*, to appear, 2013.

[18] Umberto Mosco. On the continuity of the Young-Fenchel transform. *J. Math. Anal. Appl.*, 35:518–535, 1971.

[19] Gunnar Pruessner. *Self-Organised Criticality: Theory, Models and Characterisation.* Self-organised Criticality: Theory, Models, and Characterisation. Cambridge University Press, 2012.

[20] Balázs Ráth and Bálint Tóth. Erdős-Rényi random graphs + forest fires = self-organized criticality. *Electron. J. Probab.*, 14:no. 45, 1290–1327, 2009.

[21] R. Tyrrell Rockafellar. *Convex analysis.* Princeton Mathematical Series, No. 28. Princeton University Press, 1970.

[22] Didier Sornette. *Critical phenomena in natural sciences.* Springer Series in Synergetics. Springer-Verlag, second edition, 2006. Chaos, fractals, selforganization and disorder: concepts and tools.

[23] Donald L. Turcotte. Self-organized criticality. *Reports on Progress in Physics*, 62(10):1377, 1999.