NON ASYMPTOTIC VARIANCE BOUNDS AND DEVIATION INEQUALITIES BY OPTIMAL TRANSPORT

KEVIN TANGUY
UNIVERSITY OF ANGERS, FRANCE

Abstract. The purpose of this note is to show how simple Optimal Transport arguments, on the real line, can be used in Superconcentration theory. This methodology is efficient to produce sharp non-asymptotic variance bounds for various functionals (maximum, median, l^p norms) of standard Gaussian random vectors in $\mathbb{R}^n$. The flexibility of this approach can also provide exponential deviation inequalities reflecting preceding variance bounds. As a further illustration, usual laws from Extreme theory and Coulomb gases are studied.

1. Introduction

As an introduction we recall some facts about Gaussian concentration of measure (cf. [16]) and Superconcentration theory (cf. [10]).

It is well known that concentration of measure is an effective tool in various mathematical areas (cf. [8]). In a Gaussian setting, classical concentration results typically state that, for a Lipschitz function $f : \mathbb{R}^n \to \mathbb{R}$ with Lipschitz constant $\|f\|_{\mathrm{Lip}}$,

$$\gamma_n\big(|f - \mathbb{E}_{\gamma_n}[f]| \geq t\big) \leq 2e^{-\frac{t^2}{2\|f\|^2_{Lip}}}, \quad t \geq 0, \tag{1.1}$$

with γ_n the standard Gaussian measure on $\mathbb{R}^n$. Another example of concentration of measure is the Poincaré inequality satisfied by γ_n. Namely, for $f \in L^2(\gamma_n)$ smooth enough :

$$\mathrm{Var}_{\gamma_n}(f) \leq \int_{\mathbb{R}^n} |\nabla f|^2 d\gamma_n, \tag{1.2}$$

where $|\cdot|$ stands for the Euclidean norm on $\mathbb{R}^n$. As effective as (1.1) and (1.2) are, their generality can lead to sub-optimal bounds in some particular cases. For instance, consider the 1-Lipschitz function on $\mathbb{R}^n$ $f(x) = \max_{i=1,\dots,n} x_i$. At the level of the variance, (1.2) gives

$$\mathrm{Var}(M_n) \leq 1,$$

with $M_n = \max_{i=1,\dots,n} X_i$ where $(X_1, \dots, X_n)$ stands for a standard Gaussian random vector in $\mathbb{R}^n$, whereas it has been proven that $\mathrm{Var}(M_n) \leq C/\log n$ with $C > 0$ a numerical constant. At an exponential level (1.1) is not satisfying either. Indeed, it is well known in Extreme theory (cf. [15]) that M_n can be renormalized by some numerical constants $a_n = \sqrt{2\log n}$ and $b_n = a_n - \frac{\log 4\pi + \log\log n}{2a_n}$, $n \geq 1$, such that

Date: Note of January 5, 2019.

$$a_n(M_n - b_n) \to \Lambda_0$$

in distribution, as $n \to \infty$, where Λ_0 corresponds to the Gumbel distribution :

$$\mathbb{P}(\Lambda_0 \leq x) = \exp(-e^{-x}), \quad x \in \mathbb{R}.$$

Then, it is clear that the asymptotics of Λ_0 are not Gaussian but rather exponential on the right tail and double exponential on the left tail. It is now obvious that (1.1) and (1.2) lead to sub-optimal results for the function $f(x) = \max_{i=1,\ldots,n} x_i$. This is referred to as Superconcentration phenomenon (cf. [10]) This kind of phenomenon occurs for different functionals of Gaussian random variables (and also, as we will see, for other laws of probability) and has been studied in [7, 22, 23, 17, 24]. . . .

The purpose of this note is to show how simple transport arguments on the real line can easily lead to weighted Poincaré inequalities together with deviation inequalities which are relevant in Superconcentration theory. In particular, we will emphasize the fact that such results can be obtained by transporting the Exponential measure toward the measure of interest.

Let us describe the setting of our work before stating our main results. Let μ and ν be two probability measures on $\mathbb{R}$. Assume that both of these measures are absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}$. More precisely, assume that there exist two smooth functions $g : \mathbb{R} \to \mathbb{R}$ and $h : \mathbb{R} \to \mathbb{R}$ such that

$$d\mu(x) = h(x)dx, \quad d\nu(x) = g(x)dx$$

Then, let X be a random variable with law μ and Y be a random variable with law ν. Denote by H (respectively by G) the cumulative distribution function of X (respectively Y) and define the hazard function associated to the probability measure μ by

$$\kappa_\mu(x) = \frac{h(x)}{1 - H(x)}, \quad x \in \mathrm{supp}(\mu) \subset \mathbb{R}.$$

Similarly, κ_ν will be the hazard function associated to ν.

We also assume that ν satisfies a Poincaré inequality on $\mathbb{R}$ with constant $C_\nu > 0$. That is to say, for $f : \mathbb{R} \to \mathbb{R}$ smooth enough,

$$\mathrm{Var}_\nu(f) \leq C_\nu \int_{\mathbb{R}} f'^2 d\nu.$$

Remark. It is known (cf. [16]) that $\nu^n = \nu \otimes \ldots \otimes \nu$ also satisfies a Poincaré inequality with the same constant C_ν.

We denote by $T : \mathbb{R}^n \to \mathbb{R}^n$ the transport map between μ^n and ν^n. It satisfies, for any Borelian function $f : \mathbb{R}^n \to \mathbb{R}$,

$$\mathbb{E}_{\mu^n}(f) = \mathbb{E}_{\nu^n}\big(f \circ T\big)$$

where $T(x_1, \ldots, x_n) = \big(t(x_1), \ldots, t(x_n)\big)$ and $t : \mathbb{R} \to \mathbb{R}$ is the monotone rearrangement map pushing ν toward μ (cf. section two).

In the sequel of this note (unless stated otherwise), $Y = (Y_1, \ldots, Y_n)$ will stand for a random vector in $\mathbb{R}^n$ with $\mathcal{L}(Y) = \nu^n$ and $X = (X_1, \ldots, X_n)$ for a random vector in $\mathbb{R}^n$ with $\mathcal{L}(X) = \mu^n$.

Now, let us state our main results.

Theorem 1.1. *With the previous notations, for any function $f : \mathbb{R}^n \to \mathbb{R}$ smooth enough, $n \geq 1$, we have*

$$\operatorname{Var}\big(f(X)\big) \leq C_\nu \sum_{i=1}^{n} \mathbb{E}\left[(\partial_i f)^2 \circ T(Y) \left(\frac{\kappa_\nu(Y_i)}{\kappa_\mu\big(t(Y_i)\big)}\right)^2\right]. \tag{1.3}$$

As we will see, Theorem 1.1 can be used to obtain an exponential deviation inequality for $M_n = \max_{i=1,\dots,n} X_i$.

Theorem 1.2. *Assume that there exists a function $x \mapsto \psi(x)$ from $\mathbb{R}$ to $\mathbb{R}$, non-increasing such that*

$$\left|\frac{\kappa_\nu\big(t^{-1}(x)\big)}{\kappa_\mu(x)}\right| \leq \psi(x), \quad x \in \mathbb{R}$$

and there exists ϵ_n such that

$$\mathbb{E}\Big[\psi(M_n)^2\Big] \leq \epsilon_n.$$

Then, for any $t \geq 0$ and $n \geq 1$,

$$\mathbb{P}(\sqrt{\epsilon_n}\big(M_n - \mathbb{E}[M_n]\big) \geq t) \leq 3e^{-t}.$$

Remark. As it will be clear in the sequel, the same arguments apply to any other order statistics obtained from the random vector X.

To ease the understanding of our results, we give below an application of them when ν is the (symmetric) Exponential measure on $\mathbb{R}$ and μ is the standard Gaussian measure γ_1 on $\mathbb{R}$.

Proposition 1.1. *For $f : \mathbb{R}^n \to \mathbb{R}$ smooth enough and $n \geq 1$, we have*

$$\operatorname{Var}_{\gamma_n}(f) \leq C \sum_{i=1}^{n} \mathbb{E}_{\gamma_n}\left[(\partial_i f)^2(X) \left(\frac{1}{1+|X_i|}\right)^2\right]$$

with $C > 0$ a numerical constant.

In particular, applied to (a smooth approximation of) $f(x) = \max_{i=1,\dots,n} x_i$, we get, for every $n \geq 1$,

$$\operatorname{Var}(M_n) \leq C\mathbb{E}\left[\frac{1}{1+M_n^2}\right] \leq \frac{C}{1+\log n} \tag{1.4}$$

Proposition 1.2. *The following deviation inequality holds, for any $n \geq 1$,*

$$\gamma_n\big(M_n - \mathbb{E}[M_n] \geq t\big) \leq 3e^{-ct\sqrt{\log n}}, \quad t \geq 0 \tag{1.5}$$

Remark. Notice that preceding results improve upon classical concentration of measure (namely (1.1) and (1.2)) and can also be used for other functionals such as the Median.

Throughout the article C will stand for a positive numerical constant which may change at each occurrence.

2. Tools and proofs of the main results

2.1. **Basic facts.** First, let us present the elementary tools from Optimal Transport, on the real line, that are needed in the sequel. We want to highlight the fact that we will mostly choose (in practice) ν as the Exponential measure on $\mathbb{R}_+$ (or as the symmetric Exponential measure on $\mathbb{R}$) from which we will improve some concentration properties satisfied by the measure of interest μ. However, we will not specify the measures μ and ν in the statement of our results.

Recall that the monotone transport from ν to μ (cf. [25] for more details) is obtained by a mapping $t : \mathbb{R} \to \mathbb{R}$ such that, for every $x \in \mathbb{R}$,

(2.1)
$$G(x) = \mathbb{P}(Y \leq x) = \int_{-\infty}^{x} d\nu = \int_{-\infty}^{t(x)} d\mu = \mathbb{P}\big(X \leq t(x)\big) = H\big(t(x)\big), \quad x \in \mathbb{R}.$$

Which leads, after differentiation, to the following equality

$$g(x) = h\big(t(x)\big)t'(x), \quad x \in \mathbb{R}. \tag{2.2}$$

Then, the map $T : \mathbb{R}^n \to \mathbb{R}^n$ defined by $T(x) = \big(t(x_1), \ldots, t(x_n)\big)$, for every $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, transports ν^n on μ^n. In particular, for any $f : \mathbb{R}^n \to \mathbb{R}$ smooth enough,

$$\mathrm{Var}_{\mu^n}(f) = \mathrm{Var}_{\nu^n}\big(f \circ T\big).$$

The following Lemma (cf. [16]) is also useful in the sequel.

Lemma 2.1. *Let X a be centered random variable such that, for any $0 < \theta < \frac{1}{2\sqrt{K_n}}$,*

$$\mathrm{Var}(e^{\theta X/2}) \leq \frac{\theta^2}{4} K_n \mathbb{E}[e^{\theta X}].$$

Then, there exists $c > 0$ such that :

$$\mathbb{P}(X \geq t\sqrt{K_n}) \leq 3e^{-ct}, \qquad \forall\ t \geq 0.$$

Remark. This Lemma has been fruitfully used in recent articles about Superconcentration (cf. [12, 11, 22]).

Lemma 2.1 will be combined with Harris's negative association inequality (cf. [8]) in order to prove the deviation inequality from Theorem 1.2.

Now, let us state Harris's result. Recall that a function $f : \mathbb{R}^n \to \mathbb{R}$ is considered to be non-increasing (respectively non-decreasing) if it is non-increasing, (respectively non-decreasing) in each coordinate while the others are fixed.

Proposition 2.1. *[Harris] Let $f : \mathbb{R}^n \to \mathbb{R}$ be a non-decreasing function and $g : \mathbb{R}^n \to \mathbb{R}$ be a non-increasing function. Let $X_1, \ldots, X_n$ be independent random variables and set $X = (X_1, \ldots, X_n)$. Then*

$$\mathbb{E}\big[f(X)g(X)\big] \leq \mathbb{E}\big[f(X)\big]\mathbb{E}\big[g(X)\big]. \tag{2.3}$$

Remark. As explained in detail later on, Harris's negative association was a crucial argument in the study of order statistics in [7].

2.2. **Variance bounds.** We give below the proof of Theorem 1.1.

Proof. Since T transports ν^n on μ^n, we have

$$\mathrm{Var}_{\mu^n}(f) = \mathrm{Var}_{\nu^n}\big(f \circ T\big).$$

Then, one can apply the Poincaré inequality satisfied by the measure ν^n, to the function $f \circ T$:

$$\mathrm{Var}_{\nu^n}(f \circ T) \leq C_P \sum_{i=1}^{n} \int_{\mathbb{R}^n} (\partial_i f)^2 \circ T(x) t'^2(x_i) d\nu^n(x).$$

Besides, relation (2.2) yields that

$$t'(x) = \frac{g(x)}{1 - G(x)} \times \frac{1 - H\big(t(x)\big)}{h\big(t(x)\big)} = \frac{\kappa_\nu(x)}{\kappa_\mu\big(t(x)\big)}, \quad x \in \mathbb{R},$$

under the condition that $h(x) > 0, x \in \mathbb{R}$.

□

Remark. As we will see on the examples, the important step will be to estimate the behavior of the transport map t in order to get some relevant bounds on the variance of various functionals.

Notice that this approach is reminiscent of some previous work of Barthe and Roberto [3] or Gozlan [14] on the so-called weighted Poincaré inequalities on the real line. Although our approach is similar in nature, the method of Barthe and Roberto relies on Hardy's inequality whereas ours is based on monotone rearrangement argument on the real line. Our methodology is very similar to Gozlan's work [14] (in his article the transport map T is denoted by ω^{-1}).

2.3. **Deviation inequality.** Now, let us prove Theorem 1.2 with the combination of Theorem 1.1 together with Lemma 2.1 and Proposition 2.1.

Recall that, given an i.i.d. sample $X_1, \ldots, X_n$ with common law μ we define M_n, $n \geq 1$, as

$$M_n = \max_{i=1,\ldots,n} X_i.$$

Theorem 1.2. For any $\theta > 0$, apply Theorem 1.1 to (a suitable approximation of) the function $e^{\theta f}$ with $f(x) = \max_{i=1,\ldots,n} x_i$. To this task, notice first that the partial derivatives $\partial_i f = 1_{A_i}$ with $A_i = \{x_i = \max_{j=1,\ldots,n} x_j\}$, for $i = 1, \ldots, n$, form a partition of $\mathbb{R}^n$ (that is to say $\sum_{i=1}^n 1_{A_i} = 1$). This yields

$$\begin{aligned} \mathrm{Var}(e^{\theta M_n/2}) &\leq C\frac{\theta^2}{4} \sum_{i=1}^{n} \mathbb{E}\Big[1_{A_i} \psi(X_i)^2 e^{\theta M_n}\Big] \\ &= C\frac{\theta^2}{4} \mathbb{E}\Big[e^{\theta M_n} \psi(M_n)^2\Big], \end{aligned}$$

with $M_n = \max_{i=1,\ldots,n} X_i$. Then, under the hypothesis of Theorem 1.2, use Harris's inequality (2.1). Thus,

$$\begin{aligned}\mathrm{Var}(e^{\theta M_n/2}) &\leq C\frac{\theta^2}{4}\mathbb{E}[e^{\theta M_n}]\mathbb{E}\big[\psi(M_n)^2\big]\\ &\leq C\frac{\theta^2}{4}\epsilon_n\mathbb{E}[e^{\theta M_n}]\end{aligned}$$

The conclusion follows easily with Lemma 2.1. □

3. Applications

In this section, we provide some applications of Theorem 1.1 and Theorem 1.2 in different mathematical areas.

3.1. **Extreme Theory.** We refer to [15, 13] for more details about Extreme Theory. Recall that, given a probability measure μ (absolutely continuous with respect to the Lebesgue measure) and an i.i.d. sample $X_1,\dots X_n$ with $\mathcal{L}(X_1)=\mu$, it is a classical fact that one can find renormalizing constants a_n and b_n such that $a_n(M_n-b_n)$ (where $M_n=\max_{i=1,\dots,n}X_i$) converges in distribution as $n\to\infty$ and the limiting distributions are now fully characterized. We show that our main results can be used to obtain non-asymptotic variance bounds and deviation inequalities in accordance to Extreme Theory.

Let us begin at the level of the variance.

3.1.1. *Non-asymptotic variance bounds.* Let us start with a pedagogical example from the Weibull's domain of attraction. To do so, we choose ν as the standard Exponential measure on $\mathbb{R}_+$ (that is to say $H(x)=1-e^{-x}$ if $x\geq 0$, $H(x)=0$ otherwise). Then, Theorem 1.1 yields the following Corollary:

Corollary 3.1. *If Y follows a standard Exponential distribution on $\mathbb{R}_+$ then, for any function $f:\mathbb{R}^n\to\mathbb{R}$ smooth enough and every $n\geq 1$,*

$$\mathrm{Var}\big(f(X)\big)\leq 4\sum_{i=1}^n\mathbb{E}\left[\left(\frac{\partial_i f(X)}{\kappa_\mu(X_i)}\right)^2\right], \tag{3.1}$$

where $X_1,\dots,X_n$ are independant random variables with distribution μ.

In particular, for (any smooth approximation of) $f(x)=\max_{i=1,\dots,n}x_i$,

$$\mathrm{Var}\big(M_n\big)\leq C\mathbb{E}\left[\left(\frac{1}{\kappa_\mu(M_n)}\right)^2\right], \tag{3.2}$$

where $M_n=\max_{i=1,\dots,n}X_i$ and $C>0$ is a numerical constant.

In particular, if μ stands for the uniform measure on $[0,1]$ we have

$$\mathrm{Var}(M_n)\leq 4\mathbb{E}[(1-M_n)^2]=O(1/n^2).$$

Proof. The first part is a straightforward application of Theorem (1.1).

Now, if μ stands for the uniform measure on $[0,1]$ we have $\kappa_\mu(x)=1_{x\in[0,1]}\frac{1}{1-x}$. Therefore,

$$\mathrm{Var}(M_n) \leq 4\mathbb{E}[(1-M_n)^2].$$

It is now an easy task to show that the previous inequality is sharp. Indeed, for any $t \in [0,1]$, $\mathbb{P}(M_n \leq t) = t^n$. This implies that the maximum M_n admits $t \mapsto nt^{n-1}1_{[0,1]}$ as density with respect to the Lebesgue measure.

Thus,

$$\mathbb{E}[M_n] = \int_0^1 nt^n dt = \frac{n}{n+1}$$

and

$$\mathbb{E}[M_n^2] = \int_0^1 nt^{n+1} dt = \frac{n}{n+2}.$$

Therefore, $\mathrm{Var}(M_n) = \frac{n}{n+1} - \frac{n^2}{(n+1)^2} = \frac{2n}{(n+2)(n+1)^2} = O(1/n^2)$. The same estimates also imply that

$$\mathbb{E}[(1-M_n)^2] = O\big(1/n^2\big).$$

□

Remark. (1) Recall that, $n(M_n - 1)$ converges in law toward the Weibull distribution. So, the preceding bound is the correct order of the variance of M_n.

(2) More generally, if μ stands for the Beta law with parameter $a, b > 0$, it is not difficult to show that, for every $x \in [0,1]$,

$$\begin{aligned} \frac{1}{\kappa_\mu(x)} &= \frac{\int_x^1 (1-t)^{b-1} t^{a-1} dt}{x^{a-1}(1-x)^{b-1}} \\ &\leq \min\left(\frac{1}{a}\frac{(1-x^a)}{x^{a-1}}, \frac{1}{b}\frac{(1-x)}{x^{a-1}}\right) \end{aligned}$$

Notice that if $a = b = 1$ we recover the estimates for the uniform measure. When $a > 0$ and $b > 0$ it seems hard to achieve the expected bound (of order n^{-b}) on the variance from the preceding estimate of κ_μ.

(3) It is also possible to send the standard exponential measure toward the Paréto distribution (which belongs to the Fréchet domain of attraction), however this leads to a trivial bound which is not really relevant.

Now, let us focus on the domain of attraction of the Gumbel distribution. To this task, we will transport the symmetric Exponential measure (on $\mathbb{R}$) ν towards strictly log-concave measures μ (on $\mathbb{R}$) (the standard Gaussian measure for instance).

Recall that ν admits the following density $g(x) = \frac{1}{2}e^{-|x|}$ with respect to the Lebesgue measure and admits $G(x) = \frac{1}{2}e^x$ if $x \leq 0$, $G(x) = 1 - \frac{1}{2}e^{-x}$ if $x > 0$ as a cumulative distribution function. Elementary calculus yields that

$$\kappa_\nu(x) = \begin{cases} 1, & x > 0, \\ \frac{1}{2e^{-x}-1}, & x \leq 0. \end{cases} \tag{3.3}$$

Thus, Theorem 1.1 implies the following Corollary:

Corollary 3.2. *If Y follows the symmetric Exponential distribution on $\mathbb{R}$ then, for any function $f : \mathbb{R}^n \to \mathbb{R}$ smooth enough,*

$$\mathrm{Var}\big(f(X)\big) \leq 4 \sum_{i=1}^n \mathbb{E}\left[\partial_i^2 f(X) \left(\frac{\kappa_\nu\big(t^{-1}(X_i)\big)}{\kappa_\mu(X_i)}\right)^2\right], \tag{3.4}$$

where $X = (X_1, \dots, X_n)$ has distribution μ^n.

Remark. Here, the constant 4 stands for the Poincaré constant of the symmetric Exponential measure (cf. [2]).

To illustrate the preceding Corollary, we will need a technical Lemma. This one is a precise estimation of the behavior of the transport map which will permit to obtain relevant bounds for the variance of the maximum of a symmetric (strictly) log-concave measure $d\mu(x) = e^{-V(x)}Z^{-1}dx$ with Z a normalizing constant (*e.g.* $V(x) = |x|^\alpha/\alpha$, $\alpha > 1$).

Lemma 3.1. *Consider the transport map t sending the symmetric of the Exponential measure ν toward the measure $d\mu(x) = e^{-V(x)}Z^{-1}dx$, where $V(x) = |x|^\alpha/\alpha$, $\alpha > 1$. Then, the following holds*

$$|t' \circ t^{-1}(x)| \leq \frac{C_\alpha}{V'(|x|) + 1}, \quad x \in \mathbb{R}$$

with $C_\alpha > 0$ a numerical constant depending only on α.

Proof. We would like to bound, for any $x \in \mathbb{R}$, the following ratio

$$t' \circ t^{-1}(x) = \frac{\kappa_\nu\big(t^{-1}(x)\big)}{\kappa_\mu(x)}, \tag{3.5}$$

with κ_ν defined by (3.3) and $\kappa_\mu(x) = e^{-V(x)}Z^{-1}\int_x^\infty e^{-V(t)}dt$, $x \in \mathbb{R}$. Recall that

$$t^{-1}(x) = G^{-1} \circ H(x),\ x \in \mathbb{R}$$

with

$$G^{-1}(y) = \begin{cases} \ln(2y), & 0 \leq y \leq 1/2, \\ \ln\left(\frac{1}{2(1-y)}\right), & 1/2 \leq y \leq 1. \end{cases} \tag{3.6}$$

Let $A > 0$ be sufficiently large. For $x > A$, the equation (3.5) is easily bounded by standard estimates (cf. [1]) and we get

$$|t' \circ t^{-1}(x)| = e^{V(x)} \int_x^\infty e^{-V(t)}dt \leq \frac{C}{V'(x)},$$

with $C > 0$.

For x belonging to the compact $[0, A]$, there exists $C > 0$ such that $|t' \circ t^{-1}(x)| \leq C$. To sum up,

$$|t' \circ t^{-1}(x)| \leq \frac{C}{V'(x)+1}, \; x > 0.$$

For $x = 0$ we have $|t' \circ t^{-1}(x)| = 1$ since $t^{-1}(0) = G^{-1} \circ H(0) = G^{-1}(1/2) = 0$ by symmetry.

Now if, $x < -A$, we get

$$|t' \circ t^{-1}(x)| \leq \frac{2e^{V(x)}}{2e^{-t^{-1}(x)} - 1},$$

since $\frac{1}{\kappa_\mu(x)} = \frac{e^{V(x)}}{\int_x^\infty e^{-V(t)}dt} \leq 2e^{V(x)}$ for $x \geq 0$.

So, it is enough to bound from above $t^{-1}(x)$ when $x < -A$ in order to conclude. Using the symmetry of the law μ, we obtain

$$\begin{aligned} t^{-1}(x) \leq \ln\big(2H(x)\big) \;&=\; \ln\Big(2\big[1 - H(-x)\big]\Big) \\ &\leq\; \ln\left[\frac{2e^{-V(-x)}}{V'(-x)}\right]. \end{aligned}$$

Thus, for $x < -A$,

$$|t' \circ t^{-1}(x)| \leq \frac{2e^{V(-x)}}{V'(-x)e^{V(-x)} - 1} \leq \frac{C}{V'(-x)}.$$

Similarly, when $-A \leq x \leq 0$, we also obtain that $|t' \circ t^{-1}(x)| \leq C$.
Finally, all of this can be rewritten as follows

$$\left|\frac{\kappa_\nu(t^{-1}(x))}{\kappa_\nu(x)}\right| \leq \frac{C}{V'(|x|)+1},$$

with $C > 0$. □

If V is the quadratic potential associated to the standard Gaussian measure, we obtain, thanks to Lemma 3.1 and Corollary 3.2, the following result (as announced in the introduction).

Proposition 3.1. *For $f : \mathbb{R}^n \to \mathbb{R}$ smooth enough, we have*

$$\mathrm{Var}_{\gamma_n}(f) \leq C \sum_{i=1}^n \mathbb{E}_{\gamma_n}\left[(\partial_i f)^2(X)\left(\frac{1}{1+|X_i|}\right)^2\right]. \tag{3.7}$$

In particular, applied to (a smooth approximation of) $f(x) = \max_{i=1,\ldots,n} x_i$, we get, for every $n \geq 1$,

$$\mathrm{Var}(M_n) \leq C\mathbb{E}\left[\frac{1}{1+M_n^2}\right] \leq \frac{C}{1+\log n} \tag{3.8}$$

Remark. Notice that inequality (3.7) has already been obtained, in dimension one, in [6, 5].

Proof. Indeed, for the function maximum, $\partial_i f = 1_{A_i}$, $i = 1, \ldots, n$ with

$$A_i = \; \{X_i = \max_{j=1,\ldots,n} X_j\}$$

and, again, observe that $(A_i)_{i=1,\ldots,n}$ is a partition of $\mathbb{R}^n$. Therefore,

$$\begin{aligned}\sum_{i=1}^n \mathbb{E}\Big[(\partial_i f)^2(X)\Big(\frac{1}{1+|X_i|}\Big)^2\Big] &\leq \mathbb{E}\Big[\frac{1}{1+M_n^2}\Big] \\ &\leq \frac{1}{1+\log n} + \mathbb{P}(M_n \leq \sqrt{\log n}) \\ &\leq \frac{1}{1+\log n} + \Big(1 - \frac{\sqrt{\log n}}{1+\log n} e^{-\log n/2}\Big)^n \\ &\leq \frac{C}{1+\log n}\end{aligned}$$

Since, for every $t \geq 0$, $\mathbb{P}(M_n \leq t) = \big(1 - \mathbb{P}(X_1 > t)\big)^n$ with X_1 a Gaussian standard random variable. Then, we can use the following estimate (cf. [17] (Lemma 2.5) or the appendix in [10]) to bound the preceding quantity : for any $t \geq 0$,

$$\mathbb{P}(X_1 > t) \geq \frac{t}{\sqrt{2\pi}(1+t^2)} e^{-t^2/2}.$$

Thus, $\mathrm{Var}(M_n) \leq \frac{C}{\log n}$. □

Remark. Let us make some remarks on what preceded.

(1) As mentioned in the introduction, $\sqrt{2\log n}(M_n - b_n)$ converges, when $n \to \infty$, in law toward the Gumbel distribution (the precise value of b_n is irrelevant here but can be found in [13, 15]). So, the previous Corollary gives a non-asymptotic variance bound of the maximum in accordance with Extreme theory. Besides, such a bound is classically obtained by hypercontractive and interpolation arguments (cf. [10]). Here, we provide an alternative proof based on Optimal Transport arguments.

(2) Let us further notice that the scheme of proof can also be performed for the function $f(x) = \mathrm{Med}(x_1, \ldots, x_n)$, $n \geq 1$,

$$\begin{aligned}\mathrm{Var}\big(\mathrm{Med}(X)\big) &\leq \frac{C}{1+n} + C\mathbb{P}(\mathrm{Med}(X) \leq \sqrt{n})^{n/2} \\ &\leq \frac{C}{1+n} + o\Big(\frac{1}{1+n}\Big) \leq \frac{C}{1+n}\end{aligned}$$

which corresponds to the correct order of magnitude of the variance of the median (cf. [7]). Notice that, as far as we know, such bounds can not be obtained by hypercontractive arguments.

More generally, if $V(x) = |x|^\alpha/\alpha, \quad \alpha > 1$, the same proof, together with the Lemma 3.1, yields

Corollary 3.3.

$$\mathrm{Var}(M_n) \leq C \sum_{i=1}^n \mathbb{E}\Big[(\partial_i f)^2(X)\Big(\frac{1}{1+V'\big(|X_i|\big)}\Big)^2\Big]$$

In particular, applied to (a smooth approximation of) $f(x) = \max_{i=1,\ldots,n} x_i$, it gives, for $n \geq N_0$ sufficiently large,

$$\mathrm{Var}(M_n) \leq C\mathbb{E}\Big[\frac{1}{V'^2(M_n)+1}\Big] \leq \frac{C}{1+C_\alpha \ln(n)^{2(\alpha-1)/\alpha}}, \tag{3.9}$$

with $C_\alpha > 0$ and $C > 0$ some numerical constants.

Proof.

$$\begin{aligned}
\mathbb{E}\left[\frac{1}{1+|M_n|^{2(\alpha-1)}}\right] &\leq \frac{1}{1+(\log n)^{2(\alpha-1)/\alpha}} + \mathbb{P}(M_n \leq (\ln n)^{1/\alpha}) \\
&\leq \frac{1}{1+(\log n)^{2(\alpha-1)/\alpha}} + [1-\mathbb{P}(X_1 \geq (\ln n)^{1/\alpha})]^n \\
&\leq \frac{1}{1+(\log n)^{2(\alpha-1)/\alpha}} + \left(1 - \frac{1}{2(\log n)^{(\alpha-1)/\alpha} n^{1/\alpha}}\right)^n \\
&\leq \frac{C}{1+C_\alpha (\log n)^{2(\alpha-1)/\alpha}}
\end{aligned}$$

Since, if X_1 stands for a random variable with law μ, we can proceed as in the Gaussian case. Indeed, $\mathbb{P}(X_1 \geq t) \sim \frac{1}{t^{\alpha-1}} e^{-t^\alpha/\alpha}$ as $t \to \infty$. In particular, for t large enough, this yields that $\mathbb{P}(X_1 \geq t) \geq \frac{1}{2t^{\alpha-1}} e^{-t^\alpha/\alpha}$. □

Remark. Following the proof (when $\alpha = 2$) of [15], it can easily be proved that

$$a_n(M_n - b_n) \to \Lambda_0,$$

in law, when $n \to \infty$, with $a_n = \sqrt{\alpha(\log n)^{2(\alpha-1)/\alpha}}$ et $b_n = (\log n)^{1/\alpha} - \frac{\log(\alpha Z) + \frac{\alpha-1}{\alpha}\log\log n}{(\log n)^{(\alpha-1)/\alpha}}$.

Therefore, Corollary 3.3 gives a non-asymptotic bound of the variance of the maximum reflecting this convergence result. We want to highlight the fact that such a bound is another example of the Superconcentration phenomenon. Nevertheless, as far as we know, such estimates can not be obtained by hypercontractive methods (when $\alpha > 2$) as in the Gaussian case.

3.1.2. *Deviation inequalities.* It is possible to use the preceding variance bounds to immediately obtain deviation inequalities thanks to Theorem 1.2.

Proposition 3.2. *The following deviation inequality holds, for any $n \geq 1$,*

$$\gamma_n\big(M_n - \mathbb{E}[M_n] \geq t\big) \leq 3e^{-ct\sqrt{\log n}}, \quad t \geq 0 \tag{3.10}$$

Remark.

1. Concerning Extreme theory, notice that this Theorem is only relevant if μ belongs to the domain of attraction of the Gumbel distribution. Indeed, the right tail of the Gumbel distribution behaves like $t \mapsto e^{-t}$ (whereas the left tail goes faster to 0 with the following asymptotic : $t \mapsto e^{-e^t}$).
2. Proposition 3.2 still holds if one substitutes γ_n with $d\mu(x) = e^{-V(x)}Z^{-1}dx$ (where $V(x) = |x|^\alpha/\alpha$, $\alpha > 1$) and uses the variance bound from (3.9) instead of the one given by (3.8).
3. Similar results can also be obtained if one replaces the maximum by another order statistics.

3.2. **Variance of $l^p, p \geq 2$ norms of standard Gaussian vector.** As a further illustration of our approach, we propose to recover some variance's bounds of l^p-norms, $p \geq 1$, of a standard Gaussian vector, obtained in [17]. The proof will be based on Proposition 3.1. We will adopt the following notation: given a vector

$x = (x_1, \dots, x_n) \in \mathbb{R}^n$ we denote by $\|x\|_p^p = \sum_{i=1}^n |x_i|^p$ its norm.

In the article of Paouris *et al.* [17], the authors have noticed that the variance of $\|X\|_p$ is not precisely estimated by classical concentration theory. More precisely, classical tools from the theory of concentration of measure such as Poincaré inequality or the isoperimetric Gaussian inequality yield the following bound

$$\mathrm{Var}(\|X\|_p) \le \max(n^{2/p-1}, 1), \quad p \ge 1.$$

According to [17], this bound is only optimal when $1 \le p \le 2$. The authors of [17] improved this bound by using precise estimates of moments of Gaussian functionnals together with logarithmic Sobolev inequality (through the so-called Talagrand's inequality). More precisely,

Theorem 3.1 (Paouris,Valettas, Zinn)**.** *Let X be a standard Gaussian vector on $\mathbb{R}^n$ then*

$$\mathrm{Var}(\|X\|_p) \le \begin{cases} C\frac{2^p}{p} n^{2/p-1}, & 2 < p \le c\log n, \\ C/\log n, & p > c\log n, \end{cases}$$

with $C, c > 0$ some numerical constants, independent of n and p.

Here, we propose to recover Theorem 3.1 with Proposition 3.1. We will only deal with the second assertion of the Proposition (the first part can be proved with similar arguments).

Proposition 3.3. *For $n \ge N_0$, we have the following inequality*

$$\mathrm{Var}(\|X\|_p) \le \frac{C}{\log n}, \quad p > c\log n,$$

with $C > 0$ a numerical constant independent of p and n.

Proof. Let $\delta > 0$ be a parameter to be chosen later on and denote by $B_\infty(0,\delta) = \{x \in \mathbb{R}^n, \, \|x\|_\infty < \delta\}$. Thus,

$$\begin{aligned}\mathrm{Var}(\|X\|_p) &\le C\sum_{i=1}^n \Big(\int_{B_\infty(0,\delta)} \frac{|x_i|^{2(p-1)}}{1+|x_i|^2} \frac{1}{\|x\|_p^{2(p-1)}} d\gamma_n(x) + \int_{B^c_\infty(0,\delta)} \frac{|x_i|^{2(p-1)}}{1+|x_i|^2} \frac{1}{\|x\|_p^{2(p-1)}} d\gamma_n(x)\Big) \\ &= C\Big(\sum_{i=1}^n \mathcal{I}_i + \mathcal{J}_i\Big)\end{aligned}$$

We recall the following relations between l^p and l^q norms, for $p < q$, which will be freely used in the sequel,

$$\|x\|_q \le \|x\|_p \le n^{1/p-1/q}\|x\|_q, \quad \forall x \in \mathbb{R}^n$$

On one hand, since $p < 2(p-1)$,

$$\begin{aligned}\sum_{i=1}^n \mathcal{I}_i &\le \int_{B_\infty(0,\delta)} \frac{\|x\|_{2(p-1)}^{2(p-1)}}{\|x\|_p^{2(p-1)}} d\gamma_n(x) \\ &\le \mathbb{P}\big(X \in B_\infty(0,\delta)\big)\end{aligned}$$

On the other hand, since $p < 2(p-2)$,

$$\begin{aligned}\sum_{i=1}^{n}\mathcal{J}_i &\leq \int_{B^c_\infty(0,\delta)} \frac{\|x\|^{2(p-2)}_{2(p-2)}}{\|x\|^{2(p-1)}_p} d\gamma_n(x) = \int_{B^c_\infty(0,\delta)} \left(\frac{\|x\|_{2(p-2)}}{\|x\|_p}\right)^{2(p-2)} \frac{1}{\|x\|_p^2} d\gamma_n(x)\\ &\leq \int_{B^c_\infty(0,\delta)} \frac{d\gamma_n(x)}{\|x\|_p^2}\\ &\leq \frac{1}{\delta^2}\mathbb{P}\big(X \in B^c_\infty(0,\delta)\big)\end{aligned}$$

Furthermore, notice that the following upper bound is satisfied

$$\begin{aligned}\mathbb{P}\big(X \in B^c_\infty(0,\delta)\big) = \mathbb{P}\big(\max_{i=1,\dots,n}|X_i| \geq \delta\big) &= \mathbb{P}(\exists j \in \{1,\dots,n\},\, |X_j| \geq \delta)\\ &\leq n\mathbb{P}(|X_1| \geq \delta) \leq 2ne^{-\delta^2/2}.\end{aligned}$$

So far we have obtained,

$$\mathrm{Var}(\|X\|_p) \leq C\left(\left[1 - \frac{\delta}{1+\delta^2}e^{-\delta^2/2}\right]^n + \frac{2ne^{-\delta^2/2}}{\delta^2}\right).$$

Then, we choose $\delta = \sqrt{2\log n}$ (with n large enough) to conclude. Indeed, we have

$$\mathbb{P}\big(X \in B_\infty(0,\delta)\big) \leq (1 - e^{-\delta^2/3})^n \sim e^{-n^{1/3}}$$

together with

$$\frac{2ne^{-\delta^2/2}}{\delta^2} = \frac{1}{\log n}.$$

In other terms

$$\mathrm{Var}(\|X\|_p) \leq C\left(o\big(\frac{1}{\log n}\big) + \frac{1}{\log n}\right) \leq \frac{C}{\log n},$$

which is the result. □

3.3. **Coulomb gazes.** This section exposes another application of our main results in another mathematical area. We want to highlight that, in this section, the factors μ_i, $i = 1,\dots,n$ (from the product measure $\mu_1 \otimes \dots \mu_n$) will not be assumed identical. This difference justifies the separation of this section from the others.

Now, let us introduce a few notions about Coulomb gazes and the results obtained by Chafaï and Péché in [9]. Let us consider a gas of charged particles $\{z_1,\dots,z_n\}$ on the complex plane $\mathbb{C}$, confined individually by the external field Q and experiencing a Coulomb pair repulsive interaction. This corresponds to the probability distribution $\mathbb{C}^n$ with density proportional to

$$(z_1,\dots,z_n) \in \mathbb{C}^n \mapsto \prod_{j=1}^{n} e^{-nQ(z_j)} \prod_{1\leq j<l\leq n} |z_j - z_k|^\beta \tag{3.11}$$

with $\beta > 0$ a fixed parameter and Q a fixed smooth function.

We will focus on the particular case where $\beta = 2$ and $Q(z) = V(|z|)$ with $V(t) = t^\alpha$, $t \geq 0$, $\alpha \geq 1$. We are interested in the study of

$$|z|_{(1)} \geq \dots \geq |z|_{(n)}, \tag{3.12}$$

the order statistics of the moduli of the Coulomb gas. Notice that $|z|_{(1)} = \max_{1\leq k\leq n}|z_k|$.

In their article, the authors proved the following representation formula

Theorem 3.2 (Chafaï-Péché). *For $\beta = 2$ and under the preceding assumptions, we have the following equality in distribution*

$$\big(|z|_{(1)},\dots,|z|_{(n)}\big) = \big(R_{(1)},\dots,R_{(n)}\big)$$

with $R_{(1)} \geq \ldots \geq R_{(n)}$ the order statistics associated to independent random variables $R_1,\dots,R_n$ where R_k, for $k = 1,\dots,n$, has a density proportional to

$$t \mapsto t^{2k-1}e^{-nV(t)}1_{t\geq 0}.$$

Remark. More precisely, the case $\beta = 2$ and $V(r) = r^2$ has been proved by Rider in [18]. Chafaï and Péche extended Rider's results when $\beta = 2$ and V satisfies some convexity assumption together with some decay conditions at infinity.

In [9], based on the representation formula, the authors also proved an asymptotic result for $|z|_{(1)}$. This is the content of next Theorem

Theorem 3.3 (Chafaï-Péché). *Let $|z|_{(1)} = \max_{1\leq k\leq n}|z_k|$ be as in (3.12), with $\beta = 2$. Suppose that $V(t) = t^\alpha$, for $t \geq 0$ and for some $\alpha \geq 1$. Set $c_n = \log n - 2\log\log n - \log 2\pi$ and*

$$a_n = 2\Big(\frac{\alpha}{2}\Big)^{1/\alpha+1/2}\sqrt{nc_n} \quad b_n = \Big(\frac{2}{\alpha}\Big)^{1/\alpha}\left(1+\frac{1}{2}\sqrt{\frac{2c_n}{\alpha n}}\right).$$

Then $\big(a_n(|z|_{(1)} - b_n)\big)_{n\geq 1}$ converges in distribution, as $n\to\infty$, toward the standard Gumbel law.

We will see that it is not difficult to get a non-asymptotic upper bound on the variance of $|z_{(1)}|$, together with a deviation inequality for our main results. A crucial step is the representation formula (3.12) of $|z_{(1)}|$:

$$|z_{(1)}| = \max_{i=1,\dots,n} R_i \quad \text{in law}$$

where $R_1,\dots,R_n$ are independent random variables and R_k, for any $k = 1,\dots,n$, has a density proportional to

$$t \mapsto t^{2k-1}e^{-nt^\alpha}1_{[0,\infty)}(t),\ \alpha \geq 1.$$

Then, it is possible to transport the standard Exponential measure on $\mathbb{R}_+^n$ toward the measure $\mu_1\otimes\cdots\otimes\mu_n$ with $\mu_k = \mathcal{L}(R_k)$ for any $k = 1,\dots,n$. Notice then, for every $k = 1,\dots,n$, that μ_k is log-concave on $\mathbb{R}_+$ with potential

$$V_k(x) = nt^\alpha - (2k-1)\log t.$$

So it is not difficult to prove (thanks to the estimates from [1]) that

$$\frac{1}{\kappa_{\mu_k}(x)} \leq \frac{C_\alpha}{nx^{\alpha-1}+1},\ x > 0$$

with $C_\alpha > 0$ a numerical constant. Thus, Proposition 3.1 yields

$$\begin{aligned}\mathrm{Var}(|z_{(1)}|) \leq \frac{C_\alpha}{n^2}\mathbb{E}\Big[\frac{1}{|z_{(1)}|^{2(\alpha-1)}}\Big] &\leq \frac{C_\alpha}{n\log n} + C_\alpha \mathbb{P}\bigg(|z_{(1)}| \leq \Big(\frac{\log n}{n}\Big)^{1/2(\alpha-1)}\bigg) \\ &\leq \frac{C_\alpha}{n\log n} + \prod_{i=1}^{n}\bigg[1-\mathbb{P}\bigg(R_i \geq \frac{\log n}{n}\Big)^{1/2(\alpha-1)}\bigg)\bigg] \\ &\leq \frac{C_\alpha}{n\log n} + o\bigg(\frac{1}{n\log n}\bigg) \\ &\leq \frac{C_\alpha}{n\log n}\end{aligned}$$

Also, Theorem 1.2 immediately gives the following deviation inequality

$$\mathbb{P}\bigg(\sqrt{n\log n}\big(|z_{(1)}| - \mathbb{E}[|z_{(1)}|\big) \geq t\bigg) \leq 6e^{-C_\alpha t},\ t \geq 0$$

where $C_\alpha > 0$ is a numerical constant that does not depend on n. In other words, we have obtained a non asymptotic deviation inequality together with a variance bound which are in accordance with Theorem 3.3. That is to say, we have proven the following result.

Proposition 3.4. *Let $\{z_1, \dots, z_n\}$ be a Coulomb gaze with density proportional to*

$$(z_1, \dots, z_n) \mapsto \prod_{j=1}^{n} e^{-nQ(z_j)} \prod_{1\leq j<k\leq n} |z_j - z_k|^2,$$

with $Q = V(|z|)$ and $V(t) = t^\alpha$, $\alpha \geq 1$. Then, for any $n > 1$, the following holds

$$\mathrm{Var}(|z_{(1)}|) \leq \frac{C_\alpha}{n\log n},$$

with $C_\alpha > 0$ a numerical constant, independent of n, and

$$\mathbb{P}\bigg(\sqrt{n\log n}\big(|z_{(1)}| - \mathbb{E}[|z_{(1)}|]\big) \geq t\bigg) \leq 3e^{-C_\alpha t},\ t \geq 0,$$

with $C_\alpha > 0$ a numerical constant independent of n.

4. Remarks and comparison with existing literature

In this section, we will briefly explain how stronger functional inequalities can be used to reach the right asymptotic of the left tail in the Gumbel's domain of attraction. Then, we will compare our main results with the existing literature.

4.1. **A few words on isoperimetric inequalities.** As we have already seen, the transport of the Exponential measure (toward a measure μ^n) permits to improve some concentration's properties of the measure μ^n. This phenomenon has already been observed by Talagrand in [21]. He used the isoperimetric inequality (involving a mixture of l^1 and l^2 balls) satisfied by the (symmetric) Exponential measure μ^n to improve the isoperimetric inequality satisfied by the standard Gaussian measure. More precisely, such an improvement can be seen on the following concentration inequality

$$\mathbb{P}\bigg(\Big|\max_{i=1,\dots,n}|X_i| - \sqrt{\log n}\Big| \geq C\frac{t}{\sqrt{\log n}}\bigg) \leq Ce^{-ct}, \quad t \geq 0 \tag{4.1}$$

Remark. (1) This type of inequality recently appeared in [22] for more general Gaussian measures.
(2) This gives the correct asymptotic behavior (with respect to Extreme Theory) of the right tail of the maximum. However, the asymptotic behavior of the left tail, in (4.1), is still sub-obtimal.

The symmetry of the (two sided) Exponential measure on $\mathbb{R}$, through Talagrand's isoperimetric inequality, seems to not make any distinctions between the left tail and the right and only gives a exponential decay. In [4], Bobkov studied a different isoperimetric problem (with the standard Exponential measure and uniform enlargements B_∞ instead). The lack of symmetry of the (standard) Exponential measure can be used to achieve the correct decay of the left tail on the maximum (in the Gumbel's domain of attraction).

More precisely, Bobkov proved the following Theorem.

Theorem 4.1 (Bobkov)**.** *Let ν^n stand for the (standard) Exponential measure on $\mathbb{R}_+$. Then, for every non empty ideal $A \subset \mathbb{R}_+^n$ such that $\nu^n(A) = \nu^n(B_\infty)$ and every $r \geq 0$, the following inequality holds :*

$$\nu^n(A + rB_\infty) \geq \nu^n(B + rB_\infty).$$

In other words,

$$\nu^n(A + rB_\infty) \geq \left[e^{-r}\big[\nu^n(A)\big]^{1/n} + (1 - e^{-r})\right]^n.$$

Remark. (1) Recall that A is an ideal of $\mathbb{R}_+^n$ if it satisfies the following condition if $x = (x_1, \dots, x_n) \in A$, $y = (y_1, \dots, y_n) \in \mathbb{R}_+^n$, $y_i \leq x_i$ for $i = 1, \dots, n$, then $y \in A$.

(2) If $n \to \infty$ and $\nu^n(A) = p$ is constant (with respect to n), the right hand side of the preceding inequality decreases and converges toward a double exponential. That is to say

$$\nu^n(A + rB_\infty) \geq \exp(-e^{-r}\log(1/p)).$$

As presented in [4], it possible to achieve the following deviations inequalities for a measure μ^n by transporting the Exponential measure ν^n.

Theorem 4.2. *(Bobkov) Let $X_1, \dots, X_n$ be i.i.d. random variables with $\mathcal{L}(X_1) = \mu \in \mathcal{F}_0$ and set $M_n = \max_{i=1,\dots,n} X_i$. Then, for every p, $0 < p < 1$, every $t \geq 0$,*

$$\mathbb{P}(M_n - m_p \geq t) \geq C\log(1/p)\exp(-ct), \tag{4.2}$$

$$\mathbb{P}(M_n - m_p < -t) \leq C\exp\big(-e^{tc}\log(1/p)\big), \tag{4.3}$$

where m_p stands for the quantile of order p of M_n and $C, c > 0$ are numerical constants.

Remark. In [4], there are some workable conditions which describe the set of measures $\mathcal{F}_0$. For instance Gamma measures or absolute value of standard Gaussian measures belong to $\mathcal{F}_0$.

In particular, if we choose p such that $p^{1/n} = F^{-1}(1-1/n)$, m_p corresponds to the renormalizing term used in Extreme theory. For instance, for the the Gamma measure, Bobkov's Theorem yields

Proposition 4.1. *Let $X_1, \dots, X_n$ be i.i.d Gamma random variables. Set $M_n = \max_{i=1,\dots,n} X_i$, then for every $t \geq 0$ and every $n \geq 1$*

$$\mathbb{P}(M_n - \log n \geq t) \leq Ce^{-ct}$$

and

$$\mathbb{P}(M_n - \log n \leq -t) \leq Ce^{-e^{ct}}$$

where $C, c > 0$ are numerical constants.

These non-asymptotic deviations inequalities express the correct tail behavior of the maximum of Gamma random variables (which belongs to the Gumbel's domain of attraction). Furthermore, such inequalities imply that $\mathbb{P}(|M_n - \log n| \geq t) \leq Ce^{-ct}$, which can be integrated to recover the fact (that can be easily obtained from Poincaré inequality) that $\mathrm{Var}(M_n) \leq C$.

All of this should be obtained for the maximum of absolutes values of independent and identically distributed standard Gaussian random variables. The details are left to the reader. Recall that this kind of inequality has already been obtained by Schechtman in [19].

4.2. **Comparison with existing literature.** In this section we compare our main results with recent articles which produce Superconcentration for i.i.d. random variables by other means.

4.2.1. *Renyi's representation and order statistics.* The authors of [7] combined three different arguments to bound the variance (or to obtain deviation inequalities) of order statistics from a sample of i.i.d. random variables. More precisely, let $X_1, \dots X_n$ be real i.i.d. random variables. Denote the associated order statistics by

$$X_{(1)} > \dots > X_{(n)}.$$

In their article [7], the authors obtained the following result

$$\mathrm{Var}(X_{(k)}) \leq \frac{2}{k}\mathbb{E}\left[\frac{1}{\kappa_\mu(X_{(k+1)})^2}\right], \quad k = 1, \dots, n.$$

Their scheme of proof is based on Renyi's representation formula (cf. [13]), which allows one to express order statistics in terms of renormalized sums of i.i.d Exponential random variables. They combined this representation with Efron-Stein's inequality (cf. [8]) and Harris's negative association (to do so they must assume that the function κ_μ is non-increasing) in order to bound from above the variance of $X_{(k)}$, $k = 1, \dots, n$.

They also obtained right deviation inequalities (around the mean) in a Gaussian setting. That is to say, if $X_i = |Y_i|$ with $\mathcal{L}(Y_i) = \mathcal{N}(0,1)$ for every $i = 1, \dots n$ and $U(s) = \Phi^{-1}(1 - 1/(2s))$, with Φ the distribution function of a standard Gaussian random variable, they obtained

$$\mathbb{P}\left(X_{(1)} - \mathbb{E}[X_{(1)}] \leq t/(3U(n) + \sqrt{t}/U(n) + \delta_n \right) \leq e^{-t},\, t \geq 0$$

with $\delta_n > 0$ and $[U(n)]^3 \delta_n \to \frac{\pi^2}{12}$ as $n \to \infty$.

The major drawback of this approach is that it can only be performed on order statistics. Our method seems to be more flexible and allows one to recover (from the measure ν) Poincaré inequality (for the measure of interest μ) when the transport map is Lipschitz. It is also clear that the non-increasing hypothesis on the function κ_μ is not necessary to obtain an upper bound on the variance. We have shown that this argument can only be used to reach exponential deviation inequalities. On this matter, Berstein's type of deviation inequality from [7] is more precise than ours, but it does not give back a relevant bound on the variance after integration. It is also surprising that the authors [7] did not deal with the more classical standard Gaussian case (without the absolute value).

4.2.2. *Hypercontractive approach and semigroup interpolations.* The comparison with the hypercontractive approach is straightforward. On one hand the hypercontractive approach can be used to deal with correlated Gaussian vectors (cf. [10, 22, 23]). On the other hand, the hypercontractive method can not reach any decay faster than $1/\log n$ and can only provide an exponential decay at the level of concentration inequalities. For instance, it does not seem possible to show, with hypercontractive arguments, that neither the variance of the Median of a standard Gaussian sample is of order $1/n$ nor to obtain the right order of the fluctuations of log-concave measure with potential $V(x) = |x|^\alpha$ when $\alpha > 2$ (notice also that hypercontractivity is not satisfied when $0 < \alpha < 1$).

4.2.3. *Comparison with Talagrand's inequality.* This section's purpose is to compare Proposition 3.1 with the following result.

Proposition 4.2 (Talagrand). *Let $f : \mathbb{R}^n \to \mathbb{R}$ be smooth enough, then:*

$$\mathrm{Var}_{\gamma_n}(f) \leq C \sum_{i=1}^{n} \frac{\|\partial_i f\|_2^2}{1 + \log\left(\frac{\|\partial_i f\|_2}{\|\partial_i f\|_1}\right)}. \tag{4.4}$$

Remark. This inequality was originally proved in [20] and has been a major tool in Superconcentration theory (cf. [10, 22, 23]).

To this task, it is enough to deal with the one dimensional case. Such inequalities are not comparable, as it can be seen on the following functions f_M and f_ϵ. Indeed, for $M > 0$ define the function f_M by

$$f_M(x) = \left(\int_0^x e^{t^2/4} 1_{[-M,M]}(t) dt \right) / \|f'_M\|_1, \quad x \in \mathbb{R}.$$

And, for every $0 < \epsilon < 1$, consider the function f_ϵ, defined by

$$f_\epsilon(x) = \begin{cases} \frac{|x|}{\epsilon} + 1, & |x| \leq \epsilon \\ 0, & |x| > \epsilon, \end{cases}$$

Then, it is enough to choose $\epsilon = 1/2n$, $n \geq 1$.

Acknowledgment. This work has been done during my Ph.D and I would like to thank my Ph.D advisor M. Ledoux for fruitful discussions. Also, I would like to *thank N. Gozlan for several comments and precious remarks. I warmly thank the referee for helpful comments in improving the exposition. This version benefited from several corrections by I. Gavra that we warmly thank for her help.*

References

[1] C. Ané, S. Blachère, D. Chafaï, P. Fougères, I. Gentil, F. Malrieu, C. Roberto, and G.Scheffer. *Sur les inégalités de Sobolev logarithmiques.* Société mathématiques de France, 2000.

[2] D. Bakry, I. Gentil, and M. Ledoux. *Analysis and geometry of Markov diffusion operators.* Grundlehren der Mathematischen Wissenschaften, 348, 2014.

[3] F. Barthe and C. Roberto. Modified logarithmic Sobolev inequalities on $\mathbb{R}$. *Potential Analysis*, 2008.

[4] S. Bobkov. Isoperimetric Inequalities for Distributions of Exponential Type. *The Annals of Probability, Vol. 22, No 2, 978-994*, 1994.

[5] S. Bobkov and C. Houdré. A converse Gaussian Poincaré-type inequality for convex functions. *Statistics and Probability Letters*, 44(3):281–290, 1999.

[6] S. Bobkov and M. Ledoux. Weighted Poincare-Type Inequalities For Cauchy And Other Convex Measures. *The Annals of Probability*, 37(2):403–427, 2009.

[7] S. Boucheron and M. Thomas. Concentration inequalities for order statistics. *Electronic Communications in Probability*, 2012.

[8] T. Boucheron, G. Lugosi, and P. Massart. *Concentration inequalities : a nonasymptotic theory of independance.* Oxford University Press, 2013.

[9] D. Chafaï and S. Péché. A note on the second order universality at the edge of Coulomb gases on the plane. *J. Stat. Phys.*, 2:368–383, 2014.

[10] S. Chatterjee. *Superconcentration and related topics.* Springer, 2014.

[11] M. Damron, J. Hanson, and P. Sosoe. Subdiffusive concentration in first-passage percolation. *Electronic Journal of Probability*, 19(109), 2014.

[12] M. Damron, J. Hanson, and P. Sosoe. Sublinear variance in first-passage percolation for general distributions. *Probab. Theory Related Fields*, 163(1-2):223–258, 2015.

[13] L. De Haan and A. Ferreira. *Extreme Value Theory.* Springer Series in Operations Research and Financial Engineering, 2006.

[14] N. Gozlan. Poincaré inequalities and dimension free concentration of measure. *Ann. Inst. Henri Poincaré Prob. Stat. 46 (2010), no. 3*, 2010.

[15] M. R. Leadbetter, G. Lindgren, and H. Rootzén. *Extremes and related properties of random sequences and processes.* Springer Series in Statistics., 1983.

[16] M. Ledoux. *The concentration of measure phenomenon.* Mathematical Surveys and Monographs, 89, 2001.

[17] G. Paouris, P. Valettas, and J. Zinn. Random version of Dvoretzky's Theorem in l_p^n. 2015.

[18] B. Rider. Order statistics and Ginibre's ensembles. *J. Stat. Phys.*, 114(3-4):1139–1148, 2004.

[19] G. Schechtman. The random version of Dvoretsky's theorem in l_∞^n. *GAFA Seminar 2004-2005*, 1910:265–270, 2007.

[20] M. Talagrand. On Russo's approximate zero-one law. *Ann. Prob. 22, 1576-1587,*, 1994.

[21] M. Talagrand. An isoperimetric theorem on the cube and the Khintchine-Kahane inequalities. *Proc. Amer. Math. Soc. 104, 905-909*, 1998.

[22] K. Tanguy. Some superconcentration inequalities for extrema of stationary gaussian processes. *Statistics and Probability Letters*, 2015.

[23] K. Tanguy. *Quelques inégalités de superconcentration : théorie et applications (in french).* PhD thesis, Institute of Mathematics of Toulouse, 2017.

[24] P. Valettas. On the tightness of Gaussian concentration for convex functions. *Journal d'Analyse Mathématique*, To appear.

[25] C. Villani. *Topics in Optimal Transportation.* AMS, 2003.