

# A Maxwell principle for generalized Orlicz balls

Samuel G. G. Johnston and Joscha Prochno

## Abstract

In [A dozen de Finetti-style results in search of a theory, Ann. Inst. H. Poincaré Probab. Statist. 23(2)(1987), 397–423], Diaconis and Freedman studied the low-dimensional projections of random vectors from the Euclidean unit sphere and the simplex in high dimensions, noting that the individual coordinates of these random vectors look like Gaussian and exponential random variables respectively. In subsequent works, Rachev and Rüschendorf and Naor and Romik unified these results by establishing a connection between  $\ell_p^N$  balls and a  $p$ -generalized Gaussian distribution. In this paper, we study similar questions in a significantly generalized and unifying setting, looking at low-dimensional projections of random vectors uniformly distributed on sets of the form

$$B_{\phi,t}^N := \left\{ (s_1, \dots, s_N) \in \mathbb{R}^N : \sum_{i=1}^N \phi(s_i) \leq tN \right\},$$

where  $\phi : \mathbb{R} \rightarrow [0, \infty]$  is a function satisfying some fairly mild conditions; in particular, we cover the case of Orlicz functions. Our method is different from both Rachev-Rüschendorf and Naor-Romik, based on a large deviation perspective in the form of quantitative versions of Cramér's theorem and the Gibbs conditioning principle, providing a natural framework beyond the  $p$ -generalized Gaussian distribution while simultaneously unraveling the role this distribution plays in relation to the geometry of  $\ell_p^N$  balls. We find that there is a critical parameter  $t_{\text{crit}}$  at which there is a phase transition in the behaviour of the low-dimensional projections: for  $t > t_{\text{crit}}$  the coordinates of random vectors sampled from  $B_{\phi,t}^N$  behave like uniform random variables, but for  $t \leq t_{\text{crit}}$  however the Gibbs conditioning principle comes into play, and here there is a parameter  $\beta_t > 0$  (the inverse temperature) such that the coordinates are approximately distributed according to a density proportional to  $e^{-\beta_t \phi(s)}$ .

**Keywords.** Generalized Orlicz balls, Gibbs conditioning principle, Gibbs measures, Maxwell principle, low-dimensional projections, quantitative Cramér theorem

**MSC.** Primary 60F05; Secondary 52A20, 60F10.

## 1 Introduction

### 1.1 The Maxwell principle

Over a century ago, Borel [2, Chapter 5] observed, independently of Maxwell, that if one chooses a random vector uniformly from the sphere in dimension  $N$ , then when  $N$  is large any given coordinate of the random vector is approximately Gaussian distributed. This result is commonly known as the Maxwell-Borel lemma or the Maxwell principle. More precisely, consider the Euclidean sphere

$$S_2^{N-1} := \left\{ (s_1, \dots, s_N) \in \mathbb{R}^N : \sum_{i=1}^N s_i^2 = N \right\}$$

of radius  $\sqrt{N}$  in  $\mathbb{R}^N$ , with the normalization here taken to ensure that the typical coordinate of an element of  $S_2^{N-1}$  has unit order when  $N$  is large. Suppose now  $(\Theta_1, \dots, \Theta_N)$  is a random vector chosen according to  $\sigma^N$ , the unique rotationally invariant probability measure on  $S_2^{N-1}$ . For  $k < N$ , let

$\sigma^{N \rightarrow k}$  denote the probability density function on  $\mathbb{R}^k$  associated with the marginal law of the first  $k$  coordinates  $(\Theta_1, \dots, \Theta_k)$  of  $(\Theta_1, \dots, \Theta_N)$ , which is given by

$$\sigma^{N \rightarrow k}(s_1, \dots, s_k) := \frac{\Gamma(\frac{N}{2})}{\Gamma(\frac{N-k}{2})(\pi N)^{k/2}} \left(1 - \frac{\sum_{j=1}^k s_j^2}{N}\right)^{\frac{N-k-1}{2}} \mathbb{1}_{\{s_1^2 + \dots + s_k^2 \leq N\}}.$$

It turns out that when  $N$  is large and  $k$  is small compared to  $N$ , the probability density  $\sigma^{N \rightarrow k}$  on  $\mathbb{R}^k$  is very close to the  $k$ -dimensional product  $\gamma^{\otimes k}$  of the standard Gaussian density

$$\gamma(s) := \frac{1}{\sqrt{2\pi}} e^{-s^2/2}, \quad s \in \mathbb{R}.$$

(Here and throughout we write  $\nu^{\otimes k}$  for the product measure on  $\mathbb{R}^k$  associated with a measure  $\nu$  on  $\mathbb{R}$ .) More specifically, in [5] Diaconis and Freedman supply the explicit bound

$$\int_{\mathbb{R}^k} |\sigma^{N \rightarrow k}(s) - \gamma^{\otimes k}(s)| ds \leq \frac{2(k+3)}{N-k-3}$$

on the total variation distance between the two probability density functions.

Diaconis and Freedman go on to observe seemingly connected phenomena in different settings. Indeed, in place of  $S_2^{N-1}$  take instead the simplex

$$D^{N-1} := \left\{ (s_1, \dots, s_N) \in \mathbb{R}^N : s_i \geq 0, \sum_{i=1}^N s_i = N \right\}, \quad (1)$$

and this time for  $k < N$  let  $\mu^{N \rightarrow k}$  denote the probability density function of the first  $k$  coordinates of a random vector selected uniformly from  $D^{N-1}$ . Diaconis and Freedman also give an explicit bound on the total variation distance between  $\mu^{N \rightarrow k}$  and the  $k$ -dimensional product of the standard exponential density  $\rho(s) := \mathbb{1}_{[0, \infty)}(s)e^{-s}$ , though proclaim in their paper to not have the right general theorem unifying these different but seemingly related observations.

A few years later, Rachev and Rüschemdorf [17] connected these phenomena in the setting of the  $\ell_p^N$  spheres (we write  $p$ -spheres)

$$S_p^{N-1} := \left\{ (s_1, \dots, s_N) \in \mathbb{R}^N : \sum_{i=1}^N |s_i|^p = N \right\}, \quad p \geq 1,$$

establishing an  $\ell_p$ -version of the classical Maxwell-Borel lemma. Indeed, to outline this connection here, consider the standard  $p$ -Gaussian density

$$\gamma_p(s) := \frac{1}{2p^{1/p}\Gamma(1+1/p)} \exp(-|s|^p/p), \quad s \in \mathbb{R}.$$

Rachev and Rüschemdorf study random vectors in  $\mathbb{R}^N$  distributed according to the *cone measure*  $\mu_p^N$  on  $S_p^{N-1}$ , which may be constructed as follows. If  $\zeta$  is distributed according to  $\gamma_p^{\otimes N}$ , then the normalised vector  $\zeta/||\zeta||_p$  takes values in the  $p$ -sphere and is distributed according to  $\mu_p^N$ . Rachev and Rüschemdorf use this probabilistic construction to analyze the marginal density  $\mu_p^{N \rightarrow k}$  of the first  $k$  coordinates  $(X_1, \dots, X_k)$  of a random vector  $(X_1, \dots, X_N)$  distributed according to the cone measure  $\mu_p^N$  (see, e.g., [14, Proof of Lemma 4] for a representation of  $\mu_p^{N \rightarrow k}$ ). Indeed, they show that when  $N \rightarrow \infty$  with  $k = o(N)$ , we have the total variation estimate

$$\int_{\mathbb{R}^k} |\mu_p^{N \rightarrow k}(s) - \gamma_p^{\otimes k}(s)| ds = \sqrt{\frac{2}{\pi e}} \frac{k}{N} + o(k/N). \quad (2)$$

We warn the reader that our definition of total variation is twice that of Rachev and Rüschendorf.

It transpires that the cone measure and surface measure coincide precisely when  $p \in \{1, 2, \infty\}$ , and the result of Rachev and Rüschendorf [17] beautifully connects the previously disparate observations of Diaconis and Freedman above: the case  $p = 2$  connects the Gaussian distribution to the Euclidean sphere  $S_2^{N-1}$ , and modulo certain symmetries, the  $p = 1$  case connects the exponential distribution to the simplex  $D^{N-1}$ . Rachev and Rüschendorf remarked that an analysis of the  $k$ -dimensional projections of a random vector from the arguably more natural surface measure  $\sigma_p^N$  on  $S_p^{N-1}$  would require a different treatment from that of the cone measure  $\mu_p^N$ .

With a view to tackling this problem with the surface measure  $\sigma_p^N$ , Naor and Romik [14] studied the discrepancy between the cone measure  $\mu_p^N$  and the surface measure  $\sigma_p^N$ , finding a constant  $C_p \in (0, \infty)$  such that the total variation distance between the two measures in  $N$ -dimensions may be bounded above by  $C_p / \sqrt{N}$ . They used their bound on the total variation between  $\sigma_p^N$  and  $\mu_p^N$  in conjunction with Rachev and Rüschendorf's bound (2) to show that we have the total variation bound

$$\int_{\mathbb{R}^k} |\sigma_p^{N-k}(s) - \gamma_p^{\otimes k}(s)| ds \leq C_p \left( \frac{k}{N} + \frac{1}{\sqrt{N}} \right)$$

on the  $k$ -dimensional projections of  $\sigma_p^N$  and the  $k$ -dimensional product of  $p$ -Gaussians.

## 1.2 A brief statement of our main result

While the works of Rachev and Rüschendorf [17] and Naor and Romik [14] certainly provide a gratifying answer to Diaconis and Freedman's [5] appeal for a unified theory, the purpose of the present paper is to show that these phenomena may be generalized significantly further in the setting of generalized Orlicz balls

$$B_{\phi,t}^N := \left\{ (s_1, \dots, s_N) \in \mathbb{R}^N : \sum_{i=1}^N \phi(s_i) \leq tN \right\}, \quad (3)$$

where  $t > 0$  and  $\phi : \mathbb{R} \rightarrow [0, \infty]$  is a *potential*—a function satisfying some fairly mild conditions. These conditions are given in Definition 1.1 below, but let us just say here that with the restrictions imposed on  $\phi$ , our framework includes the simplex as well as the  $\ell_p^N$  balls for  $p > 0$  and the more general case of Orlicz balls (where  $\phi$  is an even and convex function such that  $\phi(0) = 0$  and  $\phi(t) > 0$  for  $t \neq 0$ ), though in general the sets  $B_{\phi,t}^N$  we consider need not be convex, symmetric, simply connected or even compact. These generalisations are made possible through a different perspective based on the large deviation theory and the Gibbs conditioning principle.

We call  $D_\phi := \{s \in \mathbb{R} : \phi(s) < \infty\}$  the *domain* of  $\phi$ . Giving a very brief outline of our main result here, we find that there is a phase transition in the behaviour of the low-dimensional projections at the critical value  $t_{\text{crit}} = t_{\text{crit}}(\phi) \in (0, \infty]$  given by

$$t_{\text{crit}} := \begin{cases} \frac{1}{|D_\phi|} \int_{D_\phi} \phi(s) ds & : |D_\phi| < \infty, \\ \infty & : |D_\phi| = \infty, \end{cases}$$

where  $|D_\phi|$  is the Lebesgue measure of  $D_\phi$ . More specifically, we have the following:

- For  $t > t_{\text{crit}}$  (so that  $|D_\phi| < \infty$ ), in high dimensions the ball  $B_{\phi,t}^N$  is volumetrically similar (in sense that their intersection carries a lot of mass) to the  $N$ -fold product  $D_\phi^N$  of the domain  $D_\phi$ , so that the density on  $\mathbb{R}^k$  associated with the marginal distribution of the first  $k$  coordinates

of a random vector uniformly distributed on  $B_{\phi,t}^N$  is close in total variation distance to the  $k$ -dimensional product of the uniform density

$$\gamma_{\phi,\text{uni}}(s) := \frac{\mathbb{1}_{D_\phi}(s)}{|D_\phi|}, \quad s \in \mathbb{R}$$

on  $D_\phi$ , where  $|D_\phi|$  denotes the Lebesgue measure of  $D_\phi$ .

- When  $t \leq t_{\text{crit}}$ , the Gibbs conditioning principle comes into play, and the density on  $\mathbb{R}^k$  associated with the distribution of the  $k$ -dimensional projections of a random vector uniformly distributed on  $B_{\phi,t}^N$  are close in total variation distance to the  $k$ -dimensional product of the Gibbs density

$$\gamma_{\phi,-\beta_t}(s) := e^{-\beta_t \phi(s)} \mathbb{1}_{D_\phi}(s) / Z(-\beta_t), \quad s \in \mathbb{R}$$

where  $Z(-\beta_t) \in (0, \infty)$  is the normalisation constant known as partition function, and  $\beta_t > 0$  is a parameter, in statistical mechanics parlance referred to as inverse temperature, chosen so that

$$\int_{-\infty}^{\infty} \phi(s) \gamma_{\phi,-\beta_t}(s) ds = t.$$

Roughly speaking, this suggests that the coordinates of balls  $B_{\phi,t}^N$  are distributed according to a density with respect to Lebesgue measure for which larger values of  $\phi$  are penalised. The exponential parameter  $\beta_t$  increases as  $t$  decreases, so that this penalisation becomes stronger as the size of the ball  $B_{\phi,t}^N$  shrinks.

That concludes our very brief outline of our main results, in the next section we provide a more complete picture, where precise statements on the behaviour of the low-dimensional projections are given in terms of total variation distance.

### 1.3 Main results

We now present our definition for the class of functions  $\phi : \mathbb{R} \rightarrow [0, \infty]$  determining the generalized Orlicz balls  $B_{\phi,t}^N$  we consider. We call such functions potentials.

**Definition 1.1.** *A measurable function  $\phi : \mathbb{R} \rightarrow [0, \infty]$  is a potential if it is piecewise differentiable, the essential infimum of  $\phi$  is zero, for  $y < \infty$  the level sets  $\{s \in \mathbb{R} : \phi(s) = y\}$  of  $\phi$  are finite, and setting  $D_\phi := \{s \in \mathbb{R} : \phi(s) < \infty\}$ , the partition function*

$$Z(\alpha) := \int_{D_\phi} e^{\alpha \phi(s)} ds \tag{4}$$

*is finite and positive for all  $\alpha$  in some non-empty interval  $(-\infty, \alpha_{\text{max}})$  where  $\alpha_{\text{max}} := \sup\{\alpha \in \mathbb{R} : Z(\alpha) < \infty\}$ .*

**Remark 1.2.** As mentioned above, we call the set  $D_\phi = \{s \in \mathbb{R} : \phi(s) < \infty\}$  the *domain* of  $\phi$ , and write  $|D_\phi| \in (0, \infty]$  for its Lebesgue measure. Whenever  $|D_\phi| = \infty$ , clearly  $Z(\alpha)$  may only exist for negative values of  $\alpha$ . We also note that the existence of the partition function  $Z(\alpha)$  for some  $\alpha$  guarantees that for each  $y < \infty$ ,  $\phi^{-1}([0, y]) := \{s \in \mathbb{R} : \phi(s) \in [0, y]\}$  has finite Lebesgue measure, so that  $B_{\phi,t}^N$  is a subset of  $\phi^{-1}([0, tN])^N$ , and hence also has finite Lebesgue measure. In particular, whenever  $\phi$  is a potential it always makes sense to say a vector is *uniformly distributed* on  $B_{\phi,t}^N$ .

The generalized Orlicz balls  $B_{\phi,t}^N$  include various sets of interest. All subsets of  $\mathbb{R}^N$  considered in Section 1.1 are boundaries of generalized Orlicz balls: by taking  $\phi(s) = |s|^p$  we recover the  $\ell_p^N$  ball of radius  $(tN)^{1/p}$ , and the function  $\phi(s) := \infty \mathbb{1}_{(-\infty,0)}(s) + s$  corresponds to the simplex  $D^{N-1}$  defined in (1). Moreover, the classical Orlicz balls also fall into this framework, being those sets corresponding to  $\phi(0) = 0$ ,  $\phi(t) > 0$  for  $t \neq 0$ , and  $\phi$  is even and convex (note that then the set of points of non-differentiability is at most countable); see [9, Equation (3)].

Let us emphasise however that we make no further restrictions on  $\phi$ , so that while the sets  $B_{\phi,t}^N$  we consider have finite  $N$ -dimensional Lebesgue measure and are invariant under permutations of the coordinate axes, as mentioned above they need not be centered, convex, symmetric, simply connected or even compact. For an example of a set with none of these properties, we invite the reader to consider the set  $B_{\phi,t}^N$  associated with the potential

$$\phi(s) := \begin{cases} s & \text{if } s \in [k, k+2^{-k}) \text{ for some } k \in \{1, 2, \dots\}, \\ \infty & \text{otherwise.} \end{cases}$$

We now define two key quantities related to  $\phi$ . The first, the *domain supremum* of  $\phi$ , to be the essential supremum of  $\phi$  on its domain, i.e.

$$t_{\text{sup}} := \inf \left\{ y \in \mathbb{R} : |\phi^{-1}((y, \infty))| = 0 \right\} \in (0, \infty],$$

where  $|A|$  denotes the Lebesgue measure of a measurable set  $A$ . It is easily verified using the fact that  $Z(\alpha) < \infty$  for some  $\alpha$  that  $|D_\phi| = \infty$  implies  $t_{\text{sup}} = \infty$ . Our second quantity, which we already gave above, is the *domain average* of  $\phi$ , given by

$$t_{\text{crit}} := \begin{cases} \frac{1}{|D_\phi|} \int_{D_\phi} \phi(s) ds & : |D_\phi| < \infty, \\ \infty & : |D_\phi| = \infty, \end{cases}$$

where we emphasise that the integral in the former case may be infinite even when  $|D_\phi| < \infty$ . Of course  $t_{\text{crit}} \leq t_{\text{sup}}$ .

Now for any  $\alpha$  such that  $Z(\alpha) < \infty$ , we may define a *Gibbs* probability density

$$\gamma_{\phi,\alpha}(s) := \frac{e^{\alpha\phi(s)} \mathbb{1}_{D_\phi}(s)}{Z(\alpha)}, \quad s \in \mathbb{R}.$$

In the case where  $D_\phi$  has finite Lebesgue measure,  $Z(0) = |D_\phi| < \infty$ , so that we may write

$$\gamma_{\phi,\text{uni}}(s) := \gamma_{\phi,0}(s) = \frac{\mathbb{1}_{D_\phi}(s)}{|D_\phi|}, \quad s \in \mathbb{R},$$

for the uniform density on the domain of  $\phi$ .

Let us remark that

$$\frac{\partial}{\partial \alpha} \log Z(\alpha) = \frac{\int_{D_\phi} \phi(s) e^{\alpha\phi(s)} ds}{\int_{D_\phi} e^{\alpha\phi(s)} ds} = \int_{-\infty}^{\infty} \phi(s) \gamma_{\phi,\alpha}(s) ds.$$

In particular,  $W(\alpha) := \frac{\partial}{\partial \alpha} \log Z(\alpha)$  is an increasing function from  $(-\infty, \alpha_{\text{max}})$  to  $(0, t_{\text{sup}})$  satisfying  $W(0) = t_{\text{crit}}$ . The following lemma, which follows from Petrov's results [15] (see also [3, Theorem 6.2]), guarantees that for every  $t \in (0, t_{\text{sup}})$  the existence of a parameter  $\alpha_t$  such that the expectation of  $\phi$  against  $\gamma_{\phi,\alpha_t}(s) ds$  is equal to  $t$ .

**Lemma 1.3.** *For each  $t \in (0, t_{\text{sup}})$  there exists a unique parameter  $\alpha_t \in (-\infty, \alpha_{\text{max}})$  such that*

$$W(\alpha_t) = \int_{-\infty}^{\infty} \phi(s) \gamma_{\phi, \alpha_t}(s) ds = t.$$

Clearly since  $W$  is increasing and satisfies  $W(0) = t_{\text{crit}}$ , whenever  $t > t_{\text{crit}}$ ,  $\alpha_t > 0$ , and  $t < t_{\text{crit}}$  implies  $\alpha_t < 0$ .

We are now ready to present our main results on the low-dimensional projections of random vectors uniformly sampled from generalized Orlicz balls. Here and below,  $\mu_{\phi, t}^{N \rightarrow k}$  denotes the marginal density of the first  $k$  coordinates  $(X_1, \dots, X_k)$  of a random vector  $(X_1, \dots, X_N)$  chosen according to the uniform measure  $\mu_{\phi, t}^N$  on the generalized Orlicz ball  $B_{\phi, t}^N$ . That is

$$\mu_{\phi, t}^{N \rightarrow k}(s_1, \dots, s_k) := \frac{1}{|B_{\phi, t}^N|} \int_{\mathbb{R}^{N-k}} \mathbb{1}_{B_{\phi, t}^N}(s_1, \dots, s_k, s_{k+1}, \dots, s_N) ds_{k+1} \dots ds_N$$

With  $t_{\text{crit}}$  and  $t_{\text{sup}}$  now defined, shortly we state our main result, Theorem A. Beforehand however, let us briefly highlight that the case  $t \geq t_{\text{sup}}$  is trivial. Here, the generalized Orlicz ball is identical to the  $N$ -fold product of the domain of  $\phi$ , that is  $B_{\phi, t}^N = D_{\phi}^N = \{(s_1, \dots, s_N) \in \mathbb{R}^N : s_i \in D_{\phi}\}$ , and hence the coordinates of  $B_{\phi, t}^N$  are independent and uniformly distributed on  $D_{\phi}$ . In particular,  $\mu_{\phi, t}^{N \rightarrow k} = \gamma_{\phi, \text{uni}}^{\otimes k}$ .

Our main statement, Theorem A (which is actually a condensed statement of more detailed theorems which we state in full in Sections 5 and 6), concerns the low-dimensional projections of  $B_{\phi, t}^N$  in the more interesting setting in which  $t < t_{\text{sup}}$ . Here we find that a phase transition occurs at the point  $t = t_{\text{crit}}$ .

**Theorem A.** *If  $t_{\text{crit}} < t < t_{\text{sup}}$ , then the  $k$ -dimensional projections of a random vector uniformly sampled from  $B_{\phi, t}^N$  are close in distribution to a  $k$ -dimensional product of  $\gamma_{\phi, \text{uni}}$ . More specifically, there exist constants  $C = C(\phi, t)$ ,  $c = c(\phi, t) \in (0, \infty)$  depending on  $\phi$  and  $t$  such that for all  $k, N \in \mathbb{N}$  with  $k \leq N$ , we have*

$$\int_{\mathbb{R}^k} \left| \mu_{\phi, t}^{N \rightarrow k}(s) - \gamma_{\phi, \text{uni}}^{\otimes k}(s) \right| ds \leq C e^{-cN}. \quad (5)$$

*On the other hand, if  $t \leq t_{\text{crit}}$ , then the  $k$ -dimensional projections of a random vector uniformly sampled from  $B_{\phi, t}^N$  are close in distribution to a  $k$ -dimensional product of  $\gamma_{\phi, \alpha_t}$ , where  $\alpha_t$  is as in Lemma 1.3. More specifically, there exists a constant  $C = C(\phi, t) \in (0, \infty)$  depending on  $\phi$  and  $t$  such that for all  $k, N \in \mathbb{N}$  with  $k \leq N$ , we have*

$$\int_{\mathbb{R}^k} \left| \mu_{\phi, t}^{N \rightarrow k}(s) - \gamma_{\phi, \alpha_t}^{\otimes k}(s) \right| ds \leq C \frac{k}{N}. \quad (6)$$

In fact, the statement of our main results in Theorem A has been abbreviated somewhat for the sake of clarity. In Sections 5 and 6 the two cases  $t_{\text{crit}} < t$  and  $t \leq t_{\text{crit}}$  are considered separately, and in both settings we obtain fine estimates of the total variations occurring on the left-hand sides of (5) and (6). These estimates lead to the rougher bounds given in the statement of Theorem A.

## 1.4 Further discussion

Our approach to proving Theorem A is based around ideas from the theory of large deviations and statistical mechanics, specifically those centered around Cramér's theorem, the Gibbs conditioning principle, and Gibbs measures. Indeed, the framework of Gibbs measures in particular seems to be the natural one and demystifies the appearance of the  $p$ -Gaussian distribution when taking a probabilistic approach to the geometry of  $\ell_p^N$  balls. In fact, our approach, which is completely different from [14] and [17] and of independent interest, is based around a quantitative version of the Gibbs

conditioning principle that appears in a different paper by Diaconis and Freedman [6]. Somewhat surprisingly this result was neither cited in [14] nor [17] (or the recent paper [13]) even though it already contains ideas towards a unified and generalized theory in the sense of Rachev and Rüschendorf [17]. But this exact paper [6] of Diaconis and Freedman shall be the starting point for us. Let us also mention that the Gibbs conditioning principle and Gibbs measures have been successfully used to tackle other problems of a geometric flavor using probabilistic methods, e.g., [9] and [12, 13].

We now outline briefly how these approaches feature in our analysis. Take the first case  $t_{\text{crit}} < t < t_{\text{sup}}$  considered in Theorem A. The intuition here is that while  $t < t_{\text{sup}}$  ensures that  $B_{\phi,t}^N$  is a proper subset of the cube  $D_\phi^N$ , for each  $t$  of this form we have the convergence

$$\frac{|B_{\phi,t}^N|}{|D_\phi^N|} \rightarrow 1, \quad \text{as } N \rightarrow \infty \quad (7)$$

in the ratios of the Lebesgue measures of the two sets. Roughly speaking this entails that in high dimensions,  $B_{\phi,t}^N$  behaves a lot like the product set  $D_\phi^N$ , so that the marginal density of the low-dimensional projection is close in total variation to the product of the uniform density on  $D_\phi$  in the sense of (5). More specifically, we use a quantitative version of Cramer's theorem from large deviation theory to estimate the discrepancy in volume in the two sets, ultimately showing that it decays exponentially in  $N$ , leading to the bound in (5).

Now let us consider on the other hand the situation where  $t < t_{\text{crit}}$ , which we regard as the most interesting case. Here we find that in contrast to (7), either  $|D_\phi| = \infty$ , or even when  $|D_\phi| < \infty$ , we have

$$\frac{|B_{\phi,t}^N|}{|D_\phi^N|} \rightarrow 0, \quad \text{as } N \rightarrow \infty. \quad (8)$$

In any case, here we require a more delicate approach based on the Gibbs conditioning principle, for which we now provide a very rough outline. Suppose  $Y_1, Y_2, \dots$  are independent and identically distributed random variables distributed according to a probability density  $\nu$  on the real line, and suppose further that  $\mathbb{E}[Y_1] < t_0$ . The Gibbs conditioning principle is concerned with the asymptotic distribution of first  $k$  coordinates  $(Y_1, \dots, Y_k)$  conditioned on the large deviation event  $\{Y_1 + \dots + Y_N < tN\}$ . The Gibbs conditioning principle states that when  $N$  is large and  $k$  is small compared to  $N$ , then under certain conditions these first  $k$  coordinates are approximately distributed according to a  $k$ -dimensional product of the density

$$\nu_{\alpha_t}(s) := e^{\alpha_t s} \nu(s) / Z(\alpha_t), \quad s \in \mathbb{R},$$

where  $Z(\alpha_t)$  is a normalisation constant and  $\alpha_t < 0$  is a parameter chosen so that  $\int_{-\infty}^{\infty} s \nu_{\alpha_t}(s) ds$  is equal to  $t$ .

With a view to relating the Gibbs conditioning principle to our problem, in the setting where  $|D_\phi|$  is finite, we take a sequence of random variables  $X_1, X_2, \dots$  sampled uniformly and independently from  $D_\phi$ , and consider the transformed sequence  $Y_1, Y_2, \dots$  given by  $Y_i := \phi(X_i)$ . This transformation allows us to express membership of a random vector  $(X_1, \dots, X_N)$  in the subset  $B_{\phi,t}^N$  of  $D_\phi^N$  in terms of the large deviation event  $\{\phi(X_1) + \dots + \phi(X_N) \leq Nt\} = \{Y_1 + \dots + Y_N \leq Nt\}$ , and therefore use the Gibbs conditioning principle to understand the low-dimensional projections of  $(Y_1, \dots, Y_k)$ , and hence ultimately  $(X_1, \dots, X_k)$  on this event. It transpires that we are able to adapt our proof to ultimately make sense of sampling uniformly from  $D_\phi$  even when the Lebesgue measure of  $D_\phi$  is infinite, and by using a quantitative version of the Gibbs conditioning principle due to Diaconis and Freedman, we obtain the bound (6). Finally let us mention that in our treatment we regard the case  $t = t_{\text{crit}}$  as fitting into the framework of Gibbs conditioning in correspondence with the parameter  $\alpha_t = 0$ . In particular, the

convergence in total variation in this critical case happens at the linear (6) rather than exponential rate (5).

We close the introduction by taking a moment to clarify a small difference between the statement of Theorem A with the function  $\phi(s) = s^p$  and the frameworks considered by Diaconis and Freedman [5], Rachev and Rüschendorf [17], and Naor and Romik [14]. Namely, the above authors consider the boundary of the  $\ell_p^N$  ball, while we consider the interior of such sets. Ultimately we recover the same result despite the slightly different setting, namely that the low-dimensional projections are approximately  $p$ -Gaussian. Of course, let us also note that in many situations there is no difference between the results for the uniform distribution or the distribution with respect to the cone probability measure, even though sometimes different methods are required (see, e.g., [8, 10, 11, 16, 18]).

That concludes the introduction. We now take a moment to overview the remainder of the paper.

## 1.5 Overview

The remainder of the paper is structured as follows:

- In Section 2 we study the stability of certain Gibbs measures under truncations of their tail mass, proving stability of these truncations under various metrics, including large deviation and moments. Our work in this section allows us to consider potentials  $\phi$  with non-compact domains  $\{x : \phi(x) < \infty\}$  so that one cannot sample ‘uniformly’ from the domain of  $\phi$  as we have done above in the proof sketch above.
- In Section 3, we study how total variation distances are preserved under pushforwards and pullbacks, giving us the tools necessary to study potentials that are many-to-one.
- In Section 4 we supply quantitative statements of Cramer’s theorem and the Gibbs conditioning principle that are used in the following two sections to prove Theorem A.
- In Section 5 we give a proof of the Theorem A in the case  $t > t_{\text{crit}}$ .
- In Section 6 we give a proof of Theorem A in the case  $t \leq t_{\text{crit}}$ .
- In the appendix of the paper we give a proof sketch of the quantitative version of the Gibbs conditioning principle, Lemma 4.2.

## 2 Total variation and Gibbs truncations

The main task of this section is the statement and proof of Lemma 2.2 below concerning asymptotics of certain truncations of measures. First, we start with a quick lemma on the tail moments of probability measures with exponential moments.

**Lemma 2.1.** *Suppose  $f : [0, \infty) \rightarrow [0, \infty)$  is a function such that there exist constants  $c, C \in (0, \infty)$  such that*

$$f(s) \leq Ce^{-cs}.$$

*Then, for all  $\kappa \geq 0$ , and all  $L > \kappa/c$ , we have*

$$\int_L^\infty s^\kappa f(s) ds \leq \frac{CL^\kappa e^{-cL}}{c - \kappa/L} \leq C' e^{-c'L}$$

*for constants  $C', c' \in (0, \infty)$  depending on  $\kappa$  but independent of  $L$ .*

*Proof.* We have

$$\int_L^\infty s^\kappa f(s) ds \leq C \int_L^\infty s^\kappa e^{-cs} ds = CL^\kappa e^{-cL} \int_0^\infty \left(1 + \frac{s}{L}\right)^\kappa e^{-cs} ds.$$

Now use the bound  $\left(1 + \frac{s}{L}\right)^\kappa \leq e^{\kappa s/L}$ . □

Before stating Lemma 2.2, we give a brief informal statement. Suppose we have a probability measure  $\nu$  on  $(0, \infty)$  with exponential moments, and we create a truncated version  $\nu_{\langle L \rangle}$  of the measure restricted to taking values in  $[0, y]$ , but exponentially tilted so that  $\nu_{\langle L \rangle}$  has the same mean as  $\nu$ . The following lemma states that the measure  $\nu_{\langle L \rangle}$  is stable under tail truncations, in that when  $L$  is large, it has similar moments to  $\nu$  and is close in total variation distance to  $\nu$ .

**Lemma 2.2.** *Let  $\nu : [0, \infty) \rightarrow [0, \infty)$  be a probability density satisfying  $\nu(s) \leq Ce^{-cs}$  for all  $s \geq 0$  and with expectation  $\int_0^\infty s\nu(s) ds = t$ . Given  $L > t$ , we define the tilted truncation  $\nu_{\langle L \rangle}$  of  $\nu$  to be the probability density function*

$$\nu_{\langle L \rangle}(s) := \frac{e^{\alpha s} \mathbb{1}_{\{s < L\}} \nu(s)}{\int_0^L e^{\alpha s} \nu(s) ds},$$

on  $[0, \infty)$ , where  $\alpha = \alpha(L)$  is the unique parameter chosen so that  $\int_0^L s\nu_{\langle L \rangle}(s) ds = t$ .

Then there exist constants  $C', c', L_0 \in (0, \infty)$  such whenever  $L \geq L_0$ , we have the following bounds:

1. The size of  $\alpha(L)$  is bounded by  $\alpha(L) \leq C' e^{-c'L}$ .
2. The total variation between  $\nu_{\langle L \rangle}$  and  $\nu$  is bounded by

$$\|\nu_{\langle L \rangle} - \nu\| := \int_0^\infty |\nu_{\langle L \rangle}(s) - \nu(s)| ds \leq C' e^{-c'L}.$$

3. The difference between the  $j^{\text{th}}$  moment of  $\nu$  and  $\nu_{\langle L \rangle}$  is bounded by

$$\left| \int_0^\infty s^j \nu(s) ds - \int_0^\infty s^j \nu_{\langle L \rangle}(s) ds \right| \leq \frac{C'_0}{1 - j/L_0} L^{j+1} e^{-c'_0 L}.$$

*Proof.* In the case that  $\nu$  is supported on  $[0, L]$ ,  $\alpha(L) = 0$  and the measure  $\nu_{\langle L \rangle}$  is identical to  $\nu$ , so that all three statements are immediate. For the remainder of the proof we therefore assume without loss of generality that  $\nu$  is not supported on  $[0, L]$ .

We begin by showing that when  $L$  is large,  $\alpha(L)$  is small. First we note that the generating function  $G(\alpha) := \int_0^\infty e^{\alpha s} \nu(s) ds$  exists in a subset that contains  $(-\infty, c)$  and satisfies  $G(0) = 1, G'(0) = t$ , with  $G''(0) > t^2$  since  $\nu$  is non-degenerate.

Define the function

$$Q(\alpha, L) := \frac{\int_0^L s e^{\alpha s} \mathbb{1}_{\{s < L\}} \nu(ds)}{\int_0^L e^{\alpha s} \nu(ds)}. \tag{9}$$

Then since  $\nu$  is non-degenerate,  $Q(\alpha, L)$  is monotone increasing in the  $\alpha$  variable, and hence  $\alpha(L)$  is the unique solution to the equation  $Q(\alpha(L), L) = t$ . Since  $\nu$  is not supported on  $[0, L]$ , we have  $Q(\alpha, L) < t$  and therefore  $\alpha(L) > 0$ .

We now develop a lower bound for  $Q(\alpha, L)$ . Extending the range of the integral in the denominator to obtain the first inequality below, and then using the fact that  $G(\alpha) \geq 1$  whenever  $\alpha \geq 0$  to obtain the second, we have

$$Q(\alpha, L) = \frac{\int_0^\infty s e^{\alpha s} \nu(ds) - \int_L^\infty s e^{\alpha s} \nu(ds)}{\int_0^L e^{\alpha s} \nu(ds)}$$

$$\begin{aligned}
&\geq \frac{\int_0^\infty s e^{\alpha s} \nu(ds) - \int_L^\infty s e^{\alpha s} \nu(ds)}{\int_0^\infty e^{\alpha s} \nu(ds)} \\
&\geq K(\alpha) - \int_L^\infty s e^{\alpha s} \nu(ds),
\end{aligned} \tag{10}$$

where  $K(\alpha) := \frac{\partial}{\partial \alpha} \log G(\alpha)$ . With a view to bounding  $Q(\alpha, L)$  below, we now control the two quantities appearing on the right-hand-side of (10).

First we note that  $K(0) = t$ . Moreover,  $K'(0) > 0$  since  $G''(0) > G'(0)^2$ , so that in particular, there exists a  $\delta > 0$  and a  $\rho > 0$  such that for all  $\alpha \in [0, \delta]$ ,

$$K(\alpha) \geq t + \rho \alpha. \tag{11}$$

As for the integral in the final line of (10), recall that  $\nu(s) \leq C e^{-cs}$  for all  $s$ . In particular, setting  $f(s) = e^{\alpha s} \mu(s)$  in Lemma 2.1, we see that there are constants  $c_1, C_1$  such that for all  $\alpha \in [0, c/2]$ ,

$$\int_L^\infty s e^{\alpha s} \nu(ds) \leq C_1 e^{-c_1 L}. \tag{12}$$

Combining (11) and (12) in (10), we obtain

$$Q(\alpha, L) \geq t + \rho \alpha - C_1 e^{-c_1 L} \quad \text{for } \alpha \in [0, \delta \wedge c/2].$$

Now let  $L_0$  be sufficiently large so that for all  $L \geq L_0$  we have  $\frac{C_1}{\rho} e^{-c_1 L} \leq \delta \wedge c/2$ . Then plainly, for all  $L \geq L_0$ , we have

$$Q(\alpha, L) \geq t + \rho \alpha - C_1 e^{-c_1 L} \quad \text{for } \alpha \in \left[0, \frac{C_1}{\rho} e^{-c_1 L}\right]. \tag{13}$$

It follows that for all  $L \geq L_0$ ,  $Q\left(\frac{C_1}{\rho} e^{-c_1 L}, L\right) \geq t$ , and hence  $\alpha(L) \leq \frac{C_1}{\rho} e^{-c_1 L}$ , establishing the first statement of Lemma 2.2. Throughout the remainder of the proof we assume that  $L \geq L_0$ .

We now turn to proving the second point concerning the total variation distance between the measure and its tilted truncation. Since  $\nu_{\langle L \rangle}$  is supported on  $[0, L]$ , we have

$$\int_0^\infty |\nu_{\langle L \rangle}(s) - \nu(s)| ds = \int_0^L \left| \frac{e^{\alpha(L)s}}{\int_0^L e^{\alpha(L)s} \nu(s) ds} - 1 \right| \nu(s) ds + \int_L^\infty \nu(s) ds. \tag{14}$$

First we note that by Lemma 2.1 there are constants  $C_2, c_2 \in (0, \infty)$  so that the latter integral on the right-hand side of (14) may be bounded by

$$\int_L^\infty \nu(s) ds \leq C_2 e^{-c_2 L}. \tag{15}$$

We turn to bounding the former integral on the right hand side of (14). To this end, we note that since  $\alpha(L) \leq \frac{C_1}{\rho} e^{-c_1 L}$ , there are constant  $C_3, c_3 \in (0, \infty)$  such that for all  $s \in [0, L]$  we have

$$1 \leq e^{\alpha(L)s} \leq e^{\frac{C_1}{\rho} L e^{-c_1 L}} \leq 1 + C_3 e^{-c_3 L}. \tag{16}$$

On the other hand, again using Lemma 2.1 and the fact that  $e^{\alpha(L)s} \geq 1$  to obtain the lower bound below, and the upper bound in (16) to obtain the upper bound below, it may be seen that there are constants  $C_4, c_4 \in (0, \infty)$  such that

$$1 - C_4 e^{-c_4 L} \leq \int_0^L e^{\alpha(L)s} \nu(s) ds \leq 1 + C_4 e^{-c_4 L}. \tag{17}$$

Combining (16) with (17), we see that there exist constants  $C_5, c_5 \in (0, L)$  such that for every  $s \in [0, L]$ ,

$$\left| \frac{e^{\alpha(L)s}}{\int_0^L e^{\alpha(L)s} \nu(s) ds} - 1 \right| \leq C_5 e^{-c_5 L}. \quad (18)$$

Using the fact that  $\nu$  is a probability measure, (18) entails

$$\int_0^L \left| \frac{e^{\alpha(L)s}}{\int_0^L e^{\alpha(L)s} \nu(s) ds} - 1 \right| \nu(s) ds \leq C_5 e^{-c_5 L}. \quad (19)$$

In particular, using the bounds (15) and (19) in (14), we obtain the second point of the lemma.

It remains to prove the third point concerning the difference between moments of  $\nu$  and its tilted truncation  $\nu_{\langle L \rangle}$ , the proof of which follows quickly from the bound (18). Indeed, using the triangle inequality to obtain the first inequality below, and (18) to obtain the second we have

$$\begin{aligned} \left| \int_0^\infty s^j \nu(s) ds - \int_0^\infty s^j \nu_{\langle L \rangle}(s) ds \right| &\leq \int_0^L \left| \frac{e^{\alpha(L)s}}{\int_0^L e^{\alpha(L)s} \nu(s) ds} - 1 \right| s^j \nu(s) ds + \int_L^\infty s^j \nu(s) ds \\ &\leq C_5 e^{-c_5 L} \int_0^L s^j \nu(s) ds + \int_L^\infty s^j \nu(s) ds \\ &\leq C_5 L^j e^{-c_5 L} + C_6 e^{-c_6 L}, \end{aligned}$$

where the final inequality above follows from the fact that  $\nu$  is a probability measure to deal with the first term, and an application of Lemma 2.1 to handle the second. In particular there exist  $C_7, c_7 \in (0, \infty)$  such that

$$\left| \int_0^\infty s^j \nu(s) ds - \int_0^\infty s^j \nu_{\langle L \rangle}(s) ds \right| \leq C_7 e^{-c_7 L},$$

completing the proof of the third statement in Lemma 2.2.  $\square$

### 3 Total variation, pushforwards and potentials

In this section, and throughout the paper, we use the following definition.

**Definition 3.1.** *Whenever  $\pi$  is a probability density on  $\mathbb{R}^N$ , we write  $\pi^{N \rightarrow k}$  for the marginal density on  $\mathbb{R}^k$  of the first  $k$  coordinates  $(X_1, \dots, X_k)$ , where  $(X_1, \dots, X_N)$  is a random vector in  $\mathbb{R}^N$  distributed according to  $\pi(s) ds$ .*

#### 3.1 Total variation

In this section we collect several results on how the total variation metric interacts with product measures and pushforwards. While most of the tools we develop in the section are well known, we have included proofs with a view towards completeness.

When  $\pi$  and  $\lambda$  are measures on a measurable space  $(E, \mathcal{E})$ , we write

$$\|\pi - \lambda\| := 2 \sup_{A \in \mathcal{E}} |\pi(A) - \lambda(A)|$$

for the total variation distance between  $\pi$  and  $\lambda$ . Suppose that  $\pi$  and  $\lambda$  are both absolutely continuous with respect to a measure  $\nu$ , so that by the Radon-Nikodym theorem, we have  $d\pi = f d\nu$  and  $d\lambda = g d\nu$  for some measurable functions  $f, g : E \rightarrow [0, \infty]$ . Then it is easily verified that we have the alternative integral representation

$$\|\pi - \lambda\| = \int_E |f - g| d\nu \quad (20)$$

for the total variation distance between  $\pi$  and  $\lambda$ .

### 3.2 Total variation, product measures, containments and projections

The following simple lemma on the total variation distances between product measures is well known, dating back at least as far as Blum and Pathak [1]. We omit a proof.

**Lemma 3.2** (Total variation summing lemma). *Let  $\pi$  and  $\lambda$  be probability measures on  $(E, \mathcal{E})$ , and let  $\pi^{\otimes k}$  and  $\lambda^{\otimes k}$  be their respective product measures on the  $k$ -dimensional product space  $(E^k, \mathcal{E}^{\otimes k})$ . Then*

$$\|\pi^{\otimes k} - \lambda^{\otimes k}\| \leq k\|\pi - \lambda\|.$$

Suppose  $\pi$  and  $\lambda$  are probability measures on  $(E, \mathcal{E})$  and  $\mathcal{F}$  is a sub- $\sigma$ -algebra of  $\mathcal{E}$ . We write

$$\|\pi - \lambda\|_{\mathcal{F}} := 2 \sup_{F \in \mathcal{F}} |\pi(F) - \lambda(F)|.$$

Our next lemma states that if two measures on  $\mathbb{R}^N$  are close in total variation, so are their projections.

**Lemma 3.3.** *Let  $\pi$  and  $\lambda$  be probability densities on  $\mathbb{R}^N$ . Then with  $\pi^{N \rightarrow k}$  and  $\lambda^{N \rightarrow k}$  as in Definition 3.1 we have*

$$\int_{\mathbb{R}^k} \left| \pi^{N \rightarrow k}(s) - \lambda^{N \rightarrow k}(s) \right| ds \leq \int_{\mathbb{R}^N} |\pi(s) - \lambda(s)| ds.$$

*Proof.* This is a straightforward application of the triangle inequality. Indeed, for  $s = (s_1, \dots, s_k) \in \mathbb{R}^k$  and  $\zeta = (\zeta_1, \dots, \zeta_{N-k}) \in \mathbb{R}^{N-k}$ , write  $(s, \zeta) := (s_1, \dots, s_k, \zeta_1, \dots, \zeta_{N-k}) \in \mathbb{R}^N$ . Then

$$\begin{aligned} \int_{\mathbb{R}^k} \left| \pi^{N \rightarrow k}(s) - \lambda^{N \rightarrow k}(s) \right| ds &= \int_{\mathbb{R}^k} \left| \int_{\mathbb{R}^{N-k}} \pi(s, \zeta) - \lambda(s, \zeta) d\zeta \right| ds \\ &\leq \int_{\mathbb{R}^k} \int_{\mathbb{R}^{N-k}} |\pi(s, \zeta) - \lambda(s, \zeta)| d\zeta ds \\ &= \int_{\mathbb{R}^N} |\pi(s) - \lambda(s)| ds, \end{aligned}$$

which completes the proof. □

Our next lemma states that if  $A \subseteq B$ , and  $B \setminus A$  is small, then the projections of uniform random vectors from  $A$  and  $B$  are close in total variation distance.

**Lemma 3.4.** *Let  $A \subseteq B$  be measurable subsets of  $\mathbb{R}^N$  with finite Lebesgue measure, and let  $\mu_A$  and  $\mu_B$  denote the uniform densities on  $A$  and  $B$ . Then with  $\mu_A^{N \rightarrow k}$  and  $\mu_B^{N \rightarrow k}$  defined using Definition 3.1, we have*

$$\int_{\mathbb{R}^k} \left| \mu_A^{N \rightarrow k}(s) - \mu_B^{N \rightarrow k}(s) \right| ds \leq 2 \frac{|B \setminus A|}{|B|}.$$

*Proof.* Using Lemma 3.3 to obtain the inequality below, we have

$$\int_{\mathbb{R}^k} \left| \mu_A^{N \rightarrow k}(s) - \mu_B^{N \rightarrow k}(s) \right| ds \leq \int_{\mathbb{R}^N} \left| \mu_A(s) - \mu_B(s) \right| ds = 2 \frac{|B \setminus A|}{|B|},$$

completing the proof. □

### 3.3 Total variation and pushforwards

We now relate total variation distances under pushforwards of certain probability measures. Let  $(E, \mathcal{E})$  and  $(F, \mathcal{F})$  be measurable spaces. If  $\mu$  is a measure on  $E$ , and  $\Phi : E \rightarrow F$  is a measurable function, we write  $\Phi^\# \mu$  for the pushforward measure on  $F$ , defined by

$$\Phi^\# \mu(B) := \mu(\Phi^{-1}(B))$$

for measurable subsets  $B$  of  $F$ . If  $\mu$  is a probability measure on  $E$ , and  $X$  is a random variable distributed according to  $\mu$ , then  $\Phi^\# \mu$  is the law of  $\Phi(X)$ . It is straightforward to check that if  $\mu$  is absolutely continuous with respect to  $\lambda$ , then  $\Phi^\# \mu$  is absolutely continuous with respect to  $\Phi^\# \lambda$ .

The following lemma may be regarded as a more measure-theoretic formulation of Diaconis and Freedman's sufficiency lemma, [5, Lemma (2.4)], stating that total variation of certain measures is preserved under pushforwards.

**Lemma 3.5.** *Let  $E$  and  $F$  be measurable spaces, suppose  $\mu$  is a measure on  $E$  and suppose further that  $\Phi : E \rightarrow F$  is a measurable function. Suppose that  $\pi$  and  $\lambda$  are probability measures on  $E$  that are absolutely continuous with respect to  $\mu$ , and such that there exist  $f, g : F \rightarrow [0, \infty]$  such that*

$$\frac{d\pi}{d\mu} = f \circ \Phi \quad \text{and} \quad \frac{d\lambda}{d\mu} = g \circ \Phi.$$

Then

$$\|\pi - \lambda\| = \|\Phi^\# \pi - \Phi^\# \lambda\|. \quad (21)$$

*Proof.* It is easily verified that

$$\frac{d\Phi^\# \pi}{d\Phi^\# \mu} = f \quad \text{and} \quad \frac{d\Phi^\# \lambda}{d\Phi^\# \mu} = g.$$

In particular, using (20) to obtain the outer equalities below, and changing variable to obtain the central equality, we have

$$\|\Phi^\# \pi - \Phi^\# \lambda\| = \int_F |f(y) - g(y)| \Phi^\# \mu(dy) = \int_E |f(\Phi(s)) - g(\Phi(s))| \mu(ds) = \|\pi - \lambda\|$$

as required.  $\square$

### 3.4 The pushforward by a potential

We will occasionally abuse notation in the following sense: if  $\nu$  is a probability density on  $\mathbb{R}$  and  $f : \mathbb{R} \rightarrow \mathbb{R}$  is a measurable mapping, we write  $f^\# \nu$  for the probability density on  $\mathbb{R}$  associated with the pushforward by  $f$  of the measure  $\nu(s)ds$ . Now given our potential  $\phi$  and a probability measure on  $\mathbb{R}$ , we would like to understand the densities associated with pushforwards using  $\phi$ . To this end, consider the increasing function  $F : [0, \infty) \rightarrow [0, \infty)$  given by letting  $F(y)$  denote the Lebesgue measure of the set of all points  $s \in \mathbb{R}$  for which  $\phi(s) \leq y$ , that is  $F(y) := |\phi^{-1}[0, y]|$ . Suppose  $\phi$  is differentiable at  $s$  for all  $s \in \phi^{-1}(y)$ . Then it is easily verified that

$$\psi(y) := F'(y) = \sum_{s \in \phi^{-1}(y)} 1/|\phi'(s)|, \quad (22)$$

with the understanding that  $\psi(y)$  is equal to  $+\infty$  whenever there is an  $s \in \phi^{-1}(y)$  such that  $\phi'(s) = 0$ . The function  $\psi$  is defined for almost-all  $y \in [0, \infty)$ , and has the property that for all  $f$  such that  $f(\phi(s))$  is integrable,

$$\int_{-\infty}^{\infty} f(\phi(s)) ds = \int_0^{\infty} f(y) \psi(y) dy. \quad (23)$$

In particular, whenever  $\nu : \mathbb{R} \rightarrow [0, \infty)$  is a probability density of the form  $\nu(s) = f(\phi(s))$ , the pushforward  $\phi^\# \nu$  of the measure  $\nu(s)ds$  has density  $\phi^\# \nu(y) := f(y)\psi(y)$ .

Recall the partition function  $Z(\alpha) := \int_{D_\phi} e^{\alpha\phi(s)} ds$  defined in Section 1.3. We note that by (23) we may alternatively write

$$Z(\alpha) := \int_0^\infty e^{\alpha y} \psi(y) dy. \quad (24)$$

Now for all  $\alpha$  such that  $Z(\alpha) < \infty$  we may define the  $\alpha$ -tilted probability density on  $[0, \infty)$  by

$$\psi_\alpha(y) := \frac{e^{\alpha y} \psi(y)}{Z(\alpha)}. \quad (25)$$

In particular, in the setting where the Lebesgue measure of  $D_\phi$  is finite so that  $Z(0) = |D_\phi| < \infty$ , whenever  $X$  is uniformly distributed on  $D_\phi$ , the random variable  $\phi(X)$  is distributed according to the probability density

$$\psi_0(y) := \frac{\psi(y)}{|D_\phi|}. \quad (26)$$

Note that  $\psi_\alpha = \phi^\# \gamma_{\phi, \alpha}$  where  $\gamma_{\phi, \alpha}$  was defined in Section 1.3.

Finally, define the multivariate potential  $\Phi : \mathbb{R}^k \rightarrow \mathbb{R}^k$  by  $\Phi(s_1, \dots, s_k) := (\phi(s_1), \dots, \phi(s_k))$ . We note that whenever  $\nu$  is a measure on  $\mathbb{R}$ , we have

$$(\phi^\# \nu)^{\otimes k} = \Phi^\#(\nu^{\otimes k}).$$

## 4 The quantitative Cramér Theorem and Gibbs conditioning principle

In this section we provide further background on both Cramér's theorem and the Gibbs conditioning principle, ultimately giving quantitative versions of both principles that are used in the proof of Theorem A.

To this end, let  $Y_1, Y_2, \dots$  be a sequence of independent random variables identically distributed according to a probability density  $\psi$  on  $\mathbb{R}$ , and suppose that  $\mathbb{E}[Y_1] = t_0$ . Suppose further that the moment generating function  $Z(\alpha) := \int_{-\infty}^\infty e^{\alpha y} \psi(y) dy$  associated with the density exists in an open interval containing the origin. Fix  $k \in \mathbb{N}$  and let  $t > t_0$ .

According to Cramer's theorem, [4, Section 2.2], we have

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{P}(X_1 + \dots + X_N > tN) = I(t), \quad (27)$$

where  $I : [t_0, \infty) \rightarrow [0, \infty]$  is a *rate function* given by

$$I(t) := \alpha_t t - \log Z(\alpha_t),$$

where  $\alpha_t$  is the solution to  $\frac{\partial}{\partial \alpha} \log Z(\alpha) = t$ .

In the present paper we will appeal to a quantitative version of Cramér's theorem. Setting

$$\sigma_t^2 := \frac{\partial^2}{\partial \alpha^2} \log Z(\alpha)|_{\alpha=\alpha_t},$$

we have the following.

**Lemma 4.1** (Theorem 3 of Petrov [15]). *We have*

$$\mathbb{P}(Y_1 + \dots + Y_N > tN) = (1 + \varepsilon_N) \frac{1}{\sqrt{2\pi\sigma_t^2 N}} e^{-NI(t)}, \quad (28)$$

where there exists a constant  $C \in (0, \infty)$  such that  $|\varepsilon_N| \leq C/\sqrt{N}$ .

We now turn to discussing the Gibbs conditioning principle [4, Section 7.3], which asserts that as  $N \rightarrow \infty$ , conditioned on the event  $\{Y_1 + \dots + Y_N \leq tN\}$ , the  $k$ -dimensional random vector  $(Y_1, \dots, Y_k)$  converges in distribution to the  $k$ -dimensional product of the measure with density

$$\psi_{\alpha_t}(y) := e^{\alpha_t y} \psi(y) / Z(\alpha_t),$$

where  $\alpha_t < 0$  is chosen so that  $\int_{-\infty}^{\infty} \gamma_{\alpha_t}(y) y dy = t$ .

The quantitative Gibbs conditional principle, which we state shortly, gives a statement of this result in terms of total variation distances. Namely, define the constants

$$\xi_k := \frac{1}{2} \mathbb{E} \left[ \left| 1 - \frac{Z_1 + \dots + Z_k - k\mathbb{E}[Z_1]}{\sqrt{k\text{Var}[Z_1]}} \right| \right].$$

By the central limit theorem, as  $k \rightarrow \infty$ ,

$$\xi_k \rightarrow \xi := \frac{1}{2} \int_{-\infty}^{\infty} \frac{e^{-\frac{1}{2}s^2}}{\sqrt{2\pi}} |1 - s^2| ds = \sqrt{\frac{2}{\pi e}},$$

where the final equality above follows from noting  $\frac{d}{du} (ue^{-u^2/2}) = (1 - u^2)e^{-u^2/2}$ . Now for  $\theta \in (0, 1)$  define

$$Q(\theta) := \int_{-\infty}^{\infty} \left| 1 - \sqrt{1 - \theta} e^{\theta\zeta^2/2} \right| \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\zeta^2\right) d\zeta.$$

It is easily verified that  $Q'(0) = \xi$ .

We now state our quantitative version of the Gibbs conditioning principle, with a minor restatement of the form given in [6].

**Lemma 4.2** (Theorem 1.6 of [6]). *Let  $\psi_{\geq t}^{N \rightarrow k}$  be the marginal density of  $(Y_1, \dots, Y_k)$  conditioned on the event  $\{Y_1 + \dots + Y_N \geq tN\}$ . Let  $\alpha_t$  be the such that  $\frac{\partial}{\partial \alpha} \log Z(\alpha) \Big|_{\alpha=\alpha_t} = t$ . Then*

$$\int_{\mathbb{R}^k} \left| \psi_{\geq t}^{N \rightarrow k}(y) - \psi_{\alpha_t}^{\otimes k}(y) \right| dy = \begin{cases} \left(1 + \varepsilon_{k,N}^{(1)}\right) \xi_{k,\alpha_t} \frac{k}{N} & : k \text{ fixed}, N \rightarrow \infty, \\ \left(1 + \varepsilon_{k,N}^{(2)}\right) \xi \frac{k}{N} & : k = o(N), k, N \rightarrow \infty, \\ \left(1 + \varepsilon_{k,N}^{(3)}\right) Q(\theta) & : k \sim \theta N, k, N \rightarrow \infty, \end{cases}$$

where there is a universal constant  $C \in (0, \infty)$  such that setting  $C' := C \frac{\mathbb{E}[|X - \mathbb{E}[X]|^4]}{\text{Var}[X]^{3/2}}$  we have

$$|\varepsilon_{k,N}^{(1)}| \leq C' \sqrt{\frac{k}{N}}, \quad |\varepsilon_{k,N}^{(2)}|, |\varepsilon_{k,N}^{(3)}| \leq C' \left( \frac{1}{\sqrt{k}} + \frac{1}{\sqrt{N-k}} \right).$$

We now have all the tools at hand to prove Theorem A, which we do over the next two sections.

## 5 Proof of Theorem A: the $t > t_{\text{crit}}$ case

Recall that according to the first point in Theorem A, for all  $t_{\text{crit}} < t < t_{\text{sup}}$  there are constants  $C, c \in (0, \infty)$  depending on  $\phi$  and  $t$  such that for all  $k, N \in \mathbb{N}$  with  $k \leq N$ , we have

$$\int_{\mathbb{R}^k} \left| \mu_{\phi, t}^{N \rightarrow k}(s) - \gamma_{\phi, \text{uni}}^{\otimes k}(s) \right| ds \leq C e^{-cN}. \quad (29)$$

As mentioned in the introduction, we in fact prove the following stronger statement, of which the bound (29) is a consequence.

**Theorem 5.1.** *If  $t_{\text{crit}} < t < t_{\text{sup}}$ ,*

$$\int_{\mathbb{R}^k} \left| \mu_{\phi, t}^{N \rightarrow k}(s) - \gamma_{\phi, \text{uni}}^{\otimes k}(s) \right| ds = (2 + \varepsilon_{N, k}) \frac{1}{\sqrt{2\pi\sigma_t^2 N}} e^{-I(t)N},$$

where the rate function  $I: [0, t_{\text{sup}}] \rightarrow [0, \infty]$  is given by  $I(t) := t\alpha_t - \log Z(\alpha_t)$ ,  $\sigma_t^2 := \frac{\partial^2}{\partial \alpha^2} \log Z(\alpha)|_{\alpha=\alpha_t}$  and there are constants  $C = C(\phi, t)$ ,  $c = c(\phi, t) \in (0, \infty)$  such that  $|\varepsilon_{N, k}| \leq C \left( \frac{1}{\sqrt{N}} + \frac{k}{N} + e^{-c\sqrt{k}} \right)$  for all  $k, N$ .

In the remainder of this section we prove Theorem 5.1. We begin by consider the probability density  $Q_{\phi, t}^{N \rightarrow k}$  on  $\mathbb{R}^k$  given by the conditional law of the first  $k$  coordinates of a random vector uniformly distributed on  $D_\phi^N \setminus B_{\phi, t}^N$ . Namely,

$$Q_{\phi, t}^{N \rightarrow k}(s) := \frac{1}{|D_\phi|^N - |B_{\phi, t}^N|} \int_{\mathbb{R}^{N-k}} \mathbb{1}_{\{(s, \zeta) \in D_\phi^N \setminus B_{\phi, t}^N\}} d\zeta, \quad (30)$$

where for  $s = (s_1, \dots, s_k) \in \mathbb{R}^k$  and  $\zeta = (\zeta_1, \dots, \zeta_{N-k}) \in \mathbb{R}^{N-k}$ , we write  $(s, \zeta) := (s_1, \dots, s_k, \zeta_1, \dots, \zeta_{N-k}) \in \mathbb{R}^N$ . Moreover, we note that by definition  $\mu_{\phi, t}^{N \rightarrow k}$  may also be written as an integral over  $\mathbb{R}^{N-k}$ :

$$\mu_{\phi, t}^{N \rightarrow k}(s) := \frac{1}{|B_{\phi, t}^N|} \int_{\mathbb{R}^{N-k}} \mathbb{1}_{\{(s, \zeta) \in B_{\phi, t}^N\}} d\zeta. \quad (31)$$

We now work to express the total variation distance between  $\mu_{\phi, t}^{N \rightarrow k}(s)$  and  $\gamma_{\phi, \text{uni}}^{\otimes k}$  in terms of  $Q_{\phi, t}^{N \rightarrow k}$ . Indeed, by definition we have

$$\int_{\mathbb{R}^k} \left| \mu_{\phi, t}^{N \rightarrow k}(s) - \gamma_{\phi, \text{uni}}^{\otimes k}(s) \right| ds := \int_{\mathbb{R}^k} \left| \int_{\mathbb{R}^{N-k}} \left\{ \frac{\mathbb{1}_{\{(s, \zeta) \in B_{\phi, t}^N\}}}{|B_{\phi, t}^N|} - \frac{\mathbb{1}_{\{(s, \zeta) \in D_\phi^N\}}}{|D_\phi^N|} \right\} d\zeta \right| ds. \quad (32)$$

Now, for each  $s \in \mathbb{R}^k$ , since  $B_{\phi, t}^N \subseteq D_\phi^N$  we have

$$\begin{aligned} & \int_{\mathbb{R}^{N-k}} \left\{ \frac{\mathbb{1}_{\{(s, \zeta) \in B_{\phi, t}^N\}}}{|B_{\phi, t}^N|} - \frac{\mathbb{1}_{\{(s, \zeta) \in D_\phi^N\}}}{|D_\phi^N|} \right\} d\zeta \\ &= \left( \frac{1}{|B_{\phi, t}^N|} - \frac{1}{|D_\phi^N|} \right) \int_{\mathbb{R}^{N-k}} \mathbb{1}_{\{(s, \zeta) \in B_{\phi, t}^N\}} d\zeta - \frac{1}{|D_\phi^N|} \int_{\mathbb{R}^{N-k}} \mathbb{1}_{\{(s, \zeta) \in D_\phi^N \setminus B_{\phi, t}^N\}} d\zeta \\ &= \left( 1 - \frac{|B_{\phi, t}^N|}{|D_\phi^N|} \right) \left( \mu_{\phi, t}^{N \rightarrow k}(s) - Q_{\phi, t}^{N \rightarrow k}(s) \right), \end{aligned} \quad (33)$$

where we used (30) and (31) to obtain the final equality above. Plugging (33) into (32), we obtain

$$\int_{\mathbb{R}^k} \left| \mu_{\phi,t}^{N-k}(s) - \gamma_{\phi,\text{uni}}^{\otimes k}(s) \right| ds = \left( 1 - \frac{|B_{\phi,t}^N|}{|D_{\phi}|^N} \right) \int_{\mathbb{R}^k} \left| Q_{\phi,t}^{N-k}(s) - \gamma_{\phi,\text{uni}}^{\otimes k}(s) \right| ds. \quad (34)$$

The following lemma is the main part of the proof, giving a fine estimate of the integral occurring in (34).

**Lemma 5.2.** *Fix  $\theta \in (0, 1)$ . There are constants  $C = C(\theta)$ ,  $c = c(\theta) \in (0, \infty)$  such that for all integers  $k, N$  such that  $k \leq \theta N$  we have*

$$\int_{\mathbb{R}^k} \left| Q_{\phi,t}^{N-k}(s) - \gamma_{\phi,\text{uni}}^{\otimes k}(s) \right| ds = 2 - \varepsilon_{N,k},$$

where  $0 \leq \varepsilon_{N,k} \leq C \left( \frac{k}{N} + e^{-ck} \right)$ .

*Proof.* Note that both  $Q_{\phi,t}^{N-k}$  and  $\gamma_{\phi,(\alpha_t)}^{\otimes k}$  are supported on  $D_{\phi}^k$ . Moreover, consider now that the density  $Q_{\phi,t}^{N-k}$  may be written  $Q_{\phi,t}^{N-k}(s) = f(\Phi(s))$ , where  $\Phi : D_{\phi}^k \rightarrow [0, \infty)$  is given by  $\Phi(s_1, \dots, s_k) := (\phi(s_1), \dots, \phi(s_k))$  and  $f : [0, \infty)^k \rightarrow [0, \infty)$  is given by

$$f(y_1, \dots, y_k) := \frac{1}{|D_{\phi}|^N - |B_{\phi,t}^N|} \int_{\mathbb{R}^{N-k}} \mathbb{1}_{\{\phi(\zeta_1) + \dots + \phi(\zeta_{N-k}) > tN - \sum_{i=1}^k y_i\}} d\zeta_1 \dots d\zeta_{N-k}.$$

Similarly, we may also write  $\gamma_{\phi,\text{uni}}^{\otimes k}(s) := g(\Phi(s))$ , where  $g(y_1, \dots, y_k) := \frac{1}{|D_{\phi}^N|}$  (i.e. a multiple of the constant function). In particular, we are in the setting of Lemma 3.5 with  $E = D_{\phi}^k$ ,  $F = [0, \infty)^k$ , with  $\mu$  equal to the  $k$ -dimensional Lebesgue measure on  $E$ . It follows that

$$\int_{\mathbb{R}^k} \left| Q_{\phi,t}^{N-k}(s) - \gamma_{\phi,\text{uni}}^{\otimes k}(s) \right| ds = \int_{[0, \infty)^k} \left| \Phi^{\#} Q_{\phi,t}^{N-k}(y) - \Phi^{\#} \gamma_{\phi,\text{uni}}^{\otimes k}(y) \right| dy, \quad (35)$$

where  $\Phi^{\#} Q_{\phi,t}^{N-k}$  and  $\Phi^{\#} \gamma_{\phi,\text{uni}}^{\otimes k}$  denote the respective densities on  $[0, \infty)^k$  of the random vectors  $(\phi(X_1), \dots, \phi(X_k))$  and  $(\phi(Y_1), \dots, \phi(Y_k))$  where  $(X_1, \dots, X_k)$  is distributed according to density  $Q_{\phi,t}^{N-k}$  and  $(Y_1, \dots, Y_k)$  is distributed according to density  $\gamma_{\phi,\text{uni}}^{\otimes k}$ .

Now note that  $\Phi^{\#} Q_{\phi,t}^{N-k}$  is precisely the conditional density of  $(\phi(X_1), \dots, \phi(X_k))$  conditioned on the event  $\{\phi(X_1) + \dots + \phi(X_N) > tN\}$  where  $X_1, \dots, X_N$  are independent and uniformly distributed on  $D_{\phi}$ . Equivalently, by (26),  $\Phi^{\#} Q_{\phi,t}^{N-k}$  is the conditional density of  $(Y_1, \dots, Y_k)$  conditioned on the event  $\{Y_1 + \dots + Y_N > tN\}$ , where  $Y_1, \dots, Y_N$  are independent and identically distributed with density  $\psi_0(y) := \psi(y)/|D_{\phi}|$ . In particular, applying Lemma 4.2 to the variables  $Y_1, \dots, Y_k$ , and extracting a rather rough bound from Lemma 4.2, we see that there is a constant  $C = C(\phi, t) \in (0, \infty)$  such that for all  $k, N \in \mathbb{N}$  with  $k \leq N$

$$\int_{\mathbb{R}^k} \left| \Phi^{\#} Q_{\phi,t}^{N-k}(y) - \psi_{\alpha_t}^{\otimes k}(y) \right| dy \leq C \frac{k}{N} \quad (36)$$

where  $\alpha_t$  is chosen so that  $t = \frac{\partial}{\partial \alpha} \log Z(\alpha)|_{\alpha=\alpha_t}$ . (We remark that the extra precision granted by Lemma 4.2 is used more finely in the next section in our study of the case  $t \leq t_{\text{crit}}$ .)

In particular, using (35), (36) and the triangle inequality we have

$$\int_{\mathbb{R}^k} \left| Q_{\phi,t}^{N-k}(s) - \gamma_{\phi,\text{uni}}^{\otimes k}(s) \right| ds = \int_{[0, \infty)^k} \left| \Phi^{\#} \gamma_{\phi,\text{uni}}^{\otimes k}(y) - \psi_{\alpha_t}^{\otimes k}(y) \right| dy + \Delta_{k,N}, \quad (37)$$

where  $|\Delta_{k,N}| \leq Ck/N$ .

We now note that  $\Phi^\# \gamma_{\phi, \text{uni}}^{\otimes k} = \psi_0^{\otimes k}$ . In particular,

$$\int_{[0, \infty)^k} \left| \Phi^\# \gamma_{\phi, \text{uni}}^{\otimes k}(y) - \psi_{\alpha_t}^{\otimes k}(y) \right| dy = 2 \sup_{A \subseteq [0, \infty)^k} \left| \int_A \psi_{\alpha_t}^{\otimes k}(y) dy - \int_A \psi_0^{\otimes k}(y) dy \right|$$

is the total variation between the distributions of  $(Y_1, \dots, Y_k)$  and  $(Y'_1, \dots, Y'_k)$ , where the  $Y_i$  are i.i.d. with density  $\psi$ , and in particular have mean  $t_{\text{crit}}$ , and  $Y'_i$  are i.i.d. with density  $\psi_{\alpha_t}$ , and in particular have mean  $t$ . We are going to show that this total variation is nearly equal to 2 when  $k$  is large. Indeed, if we set

$$u := \frac{t_{\text{crit}} + t}{2},$$

then for large  $k$  it is likely that  $\left\{ \frac{Y_1 + \dots + Y_k}{k} \leq u \right\}$  but unlikely that  $\left\{ \frac{Y'_1 + \dots + Y'_k}{k} \leq u \right\}$ . More explicitly, setting  $A_u := \left\{ (y_1, \dots, y_k) \in [0, \infty)^k : y_1 + y_2 + \dots + y_k > uk \right\}$  and extracting a rather rough bound from the quantitative Cramér theorem, Lemma 4.1, we see that there exist constants  $C = C(\phi, t)$ ,  $c = c(\phi, t) \in (0, \infty)$  such that

$$\int_{A_u} \psi^{\otimes k}(y) dy \leq C e^{-ck} \quad \text{and} \quad \int_{A_u} \psi_{\alpha_t}^{\otimes k}(y) dy \geq 2 - C e^{-ck}.$$

In particular, two previous two estimates imply that

$$2 \geq \int_{[0, \infty)^k} \left| \Phi^\# \gamma_{\phi, \text{uni}}^{\otimes k}(y) - \psi_{\alpha_t}^{\otimes k}(y) \right| dy \geq 2 - 2C e^{-ck}. \quad (38)$$

Combining (37) with (38), we obtain the result. □

We are now ready to prove Theorem 5.1.

*Proof of Theorem 5.1.* Consider the large- $N$  asymptotics of the right hand side of (34).

On the one hand, we may write

$$\left( 1 - \frac{|B_{\phi, t}^N|}{|D_\phi|^N} \right) = \mathbb{P}[\phi(V_1) + \dots + \phi(V_N) > tN],$$

where  $V_i$  are independent random variables distributed according to  $\gamma_{\phi, \text{uni}}$ , the uniform density on  $D_\phi$ . By the quantitative version of Cramer's theorem, Lemma 4.1, we have

$$\left( 1 - \frac{|B_{\phi, t}^N|}{|D_\phi|^N} \right) = (1 + \varepsilon_N) \frac{1}{\sqrt{2\pi\sigma_t^2 N}} \exp(-I(t)N),$$

where  $|\varepsilon_N| < C/\sqrt{N}$  for a constant  $C = C(\phi, t) \in (0, \infty)$  not depending on  $N$ .

On the other hand, by Lemma 5.2 we have

$$\int_{\mathbb{R}^k} \left| Q_{\phi, t}^{N \rightarrow k}(s) - \gamma_{\phi, \text{uni}}^{\otimes k}(s) \right| ds = 2 - \varepsilon_{N, k},$$

where for a different constant  $C' = C'(\phi)$  we have  $|\varepsilon_{N, k}| \leq C' \left( \frac{k}{N} + e^{-c\sqrt{k}} \right)$ .

In particular, by (34),

$$\int_{\mathbb{R}^k} \left| \mu_{\phi, t}^{N \rightarrow k}(s) - \gamma_{\phi, \text{uni}}^{\otimes k}(s) \right| ds = (1 + \varepsilon_N)(2 - \varepsilon_{N, k}) \frac{1}{\sqrt{2\pi\sigma_t^2 N}} \exp(-I(t)N),$$

where  $|\varepsilon_{N,k}| \leq C' \left( \frac{k}{N} + e^{-c\sqrt{k}} \right)$  and  $|\varepsilon_N| < C/\sqrt{N}$ . Set  $\rho_{N,k}$  to be the solution to

$$1 + \rho_{N,k} = (1 + \varepsilon_N) \left( 1 - \frac{1}{2} \varepsilon_{N,k} \right)$$

Then plainly there are constants  $c, C \in (0, \infty)$  such that  $|\rho_{N,k}| \leq C \left( \frac{1}{\sqrt{N}} + \frac{k}{N} + e^{-ck} \right)$ , completing the proof of Theorem 5.1. □

## 6 Proof of Theorem A: the $t \leq t_{\text{crit}}$ case

### 6.1 A full statement and overview

We now turn to proving Theorem A in the case where  $t \leq t_{\text{crit}}$ . We recall from Section 1.3 that  $\alpha_t \leq 0$  is a parameter chosen so that if  $\gamma_{\phi, \alpha}(s)$  is the tilted density

$$\gamma_{\phi, \alpha}(s) := e^{\alpha \phi(s)} \mathbb{1}_{D_\phi(s)} / Z(\alpha),$$

then  $\int_{-\infty}^{\infty} \phi(s) \gamma_{\phi, \alpha}(s) ds$ .

As in the  $t > t_{\text{crit}}$ , we actually prove the following sharper result, giving a fine estimate of the total variation which implies (6).

**Theorem 6.1.** *If  $t \leq t_{\text{crit}}$ , then with  $\alpha_t$  as in Lemma 1.3 we have*

$$\int_{\mathbb{R}^k} \left| \mu_{\phi, t}^{N-k}(s) - \gamma_{\phi, \alpha_t}^{\otimes k}(s) \right| ds = \begin{cases} \left( 1 + \varepsilon_{k, N}^{(1)} \right) \xi_{k, \alpha_t} \frac{k}{N} & : k \text{ fixed}, N \rightarrow \infty, \\ \left( 1 + \varepsilon_{k, N}^{(2)} \right) \xi \frac{k}{N} & : k = o(N), k, N \rightarrow \infty, \\ \left( 1 + \varepsilon_{k, N}^{(3)} \right) Q(\theta) & : k \sim \theta N, k, N \rightarrow \infty, \end{cases}$$

and for a constant  $C = C(\phi, t) \in (0, \infty)$  we have

$$|\varepsilon_{k, N}^{(1)}| \leq C \sqrt{\frac{k}{N}}, \quad |\varepsilon_{k, N}^{(2)}|, |\varepsilon_{k, N}^{(3)}| \leq C \left( \frac{1}{\sqrt{k}} + \frac{1}{\sqrt{N-k}} \right).$$

Theorem 6.1 is proved in the remainder of this section. The proof is divided into three steps:

- Step 1. We assume that the Lebesgue measure of  $D_\phi$  is finite, and under this assumption Lemma 6.2 below estimates the total variation between the pushforwards  $\Phi^\# \mu_{\phi, t}^{N-k}$  and  $\Phi^\# \gamma_{\phi, \alpha_t}^{\otimes k}$ .
- Step 2. We will continue to assume  $|D_\phi| < \infty$ , and use our work in Section 3 to show that we may ‘pullback’ the result obtained in Lemma 6.2 to estimate the total variation between the densities  $\mu_{\phi, t}^{N-k}$  and  $\gamma_{\phi, t}^{\otimes k}$ .
- Step 3. Finally, we will show that the assumption  $|D_\phi| < \infty$  may be lifted, completing the proof of Theorem 6.1. This part is based on a truncation argument, in which the domain  $D_\phi$  of infinite Lebesgue measure is approximated by sets with finite Lebesgue measure.

The next three sections correspond to the three steps outlined above.

## 6.2 Step 1

The following lemma may be regarded as a pushforward version of Theorem 6.1

**Lemma 6.2.** *Suppose  $|D_\phi| < \infty$  and  $t \leq t_{\text{crit}}$ . Then, with  $\alpha_t \leq 0$  as in Definition 1.3, we have*

$$\int_{\mathbb{R}^k} \left| \Phi^\# \mu_{\phi,t}^{N-k}(y) - \psi_{\alpha_t}^{\otimes k}(y) \right| dy = \begin{cases} \left(1 + \varepsilon_{k,N}^{(1)}\right) \xi_{k,\alpha_t} \frac{k}{N} & : k \text{ fixed}, N \rightarrow \infty, \\ \left(1 + \varepsilon_{k,N}^{(2)}\right) \xi \frac{k}{N} & : k = o(N), k, N \rightarrow \infty, \\ \left(1 + \varepsilon_{k,N}^{(3)}\right) Q(\theta) & : k \sim \theta N, k, N \rightarrow \infty, \end{cases}$$

where there is a universal constant  $C$  with the following property. Namely, if  $X$  is a random variable distributed according to  $\psi_{\alpha_t}$ , then setting  $C' := C \frac{\mathbb{E}[(X - \mathbb{E}[X])^4]}{\text{Var}[X]^2}$  we have

$$|\varepsilon_{k,N}^{(1)}| \leq C' \sqrt{\frac{k}{N}}, \quad |\varepsilon_{k,N}^{(2)}|, |\varepsilon_{k,N}^{(3)}| \leq C' \left( \frac{1}{\sqrt{k}} + \frac{1}{\sqrt{N-k}} \right).$$

*Proof.* Since  $|D_\phi| < \infty$ , the uniform density  $\gamma_{\phi,\text{uni}}(s) := \frac{\mathbb{1}_{D_\phi}(s)}{|D_\phi|}$  on  $D_\phi$  exists. Suppose now  $X_1, \dots, X_N$  are independent random variables distributed according to the uniform density  $\gamma_{\phi,\text{uni}}$  on  $D_\phi$ , and consider the transformed variables  $\phi(Y_1), \dots, \phi(Y_N)$ , which are distributed according to  $\psi_0$ , where  $\psi_0$  is given in (26). Noting that by definition  $\mu_{\phi,t}^{N-k}$  is the marginal density of  $(X_1, \dots, X_k)$  conditioned on the event  $\{\phi(X_1) + \dots + \phi(X_N) \leq tN\}$ , it follows that the pushforward  $\Phi^\# \mu_{\phi,t}^{N-k}$  is the marginal density of  $(Y_1, \dots, Y_k)$  conditioned on the event  $\{Y_1 + \dots + Y_N \leq tN\}$ . In other words, we are in the setting of Lemma 4.2, from which the result follows immediately.  $\square$

## 6.3 Step 2

The next lemma ‘pulls back’ the previous result, replacing the estimate of the total variation between  $\Phi^\# \mu_{\phi,t}^{N-k}$  and  $\psi_{\alpha_t}^{\otimes k}$  with one between  $\mu_{\phi,t}^{N-k}$  and  $\gamma_{\alpha_t}^{\otimes k}$ . This next result amounts to precisely the statement of Theorem 6.1 in the case where the Lebesgue measure of  $D_\phi$  is finite.

**Lemma 6.3.** *Suppose  $|D_\phi| < \infty$  and  $t \leq t_{\text{crit}}$ . Then, with  $\alpha_t \leq 0$  as in Definition 1.3, we have*

$$\int_{\mathbb{R}^k} \left| \mu_{\phi,t}^{N-k}(s) - \gamma_{\alpha_t}^{\otimes k}(s) \right| ds = \begin{cases} \left(1 + \varepsilon_{k,N}^{(1)}\right) \xi_{k,\alpha_t} \frac{k}{N} & : k \text{ fixed}, N \rightarrow \infty, \\ \left(1 + \varepsilon_{k,N}^{(2)}\right) \xi \frac{k}{N} & : k = o(N), k, N \rightarrow \infty, \\ \left(1 + \varepsilon_{k,N}^{(3)}\right) Q(\theta) & : k \sim \theta N, k, N \rightarrow \infty, \end{cases}$$

where there is a universal constant  $C$  with the following property. Namely, if  $X$  is a random variable distributed according to  $\psi_{\alpha_t}$ , then setting  $C' := C \frac{\mathbb{E}[(X - \mathbb{E}[X])^4]}{\text{Var}[X]^2}$  we have

$$|\varepsilon_{k,N}^{(1)}| \leq C' \sqrt{\frac{k}{N}}, \quad |\varepsilon_{k,N}^{(2)}|, |\varepsilon_{k,N}^{(3)}| \leq C' \left( \frac{1}{\sqrt{k}} + \frac{1}{\sqrt{N-k}} \right).$$

*Proof.* We would like to use Lemma 3.5, with  $\Phi : \mathbb{R}^k \rightarrow [0, \infty)^k$  defined as in Section 3.4: i.e.  $\Phi(s_1, \dots, s_k) = (\phi(s_1), \dots, \phi(s_k))$ . To this end note that we may write  $\mu_{\phi,t}^{N-k}(s) = f(\Phi(s))$ , where

$$f(y_1, \dots, y_k) := \frac{1}{|B_{\phi,t}^N|} \int_{\mathbb{R}^{N-k}} \mathbb{1}_{\{\phi(\zeta_1) + \dots + \phi(\zeta_{N-k}) \leq tN - \sum_{i=1}^k y_i\}} d\zeta_1 \dots d\zeta_{N-k}.$$

Moreover, clearly  $\gamma_{\phi,\alpha_t}^{\otimes k}(s) = g(\Phi(s))$ , where  $g(y_1, \dots, y_k) = \frac{e^{\alpha_t(y_1 + \dots + y_k)}}{Z(\alpha)^k}$ . In particular, we are in the setting of Lemma 3.5, so that

$$\int_{\mathbb{R}^k} \left| \mu_{\phi,t}^{N-k}(s) - \gamma_{\phi,\alpha_t}^{\otimes k}(s) \right| ds = \int_{\mathbb{R}^k} \left| \Phi^\# \mu_{\phi,t}^{N-k}(y) - \Phi^\# \gamma_{\phi,\alpha_t}^{\otimes k}(y) \right| dy.$$

The result follows by noting that  $\Phi^\# \gamma_{\phi,\alpha_t}^{\otimes k} = \psi_{\alpha_t}^{\otimes k}$ , and using Lemma 6.2.  $\square$

### 6.4 Step 3

In the previous section, we proved Lemma 6.3, which is precisely the statement that Theorem 6.1 holds whenever  $D_\phi$  has finite Lebesgue measure. Here we will show that this assumption may be lifted, thereby completing the proof of Theorem 6.1.

*Proof of Theorem 6.1.* We approximate an arbitrary potential  $\phi$  by a potential  $\phi_L$  whose domain has finite Lebesgue measure. Namely, for  $L > 0$  let  $\phi_L(s) := \phi(s) + \infty \mathbb{1}\{\phi(s) > L\}$ , so that in particular  $\phi$  agrees with  $\phi_L$  on the set  $\{s \in \mathbb{R} : \phi(s) \leq L\}$ .

Define the truncated partition function

$$Z_L(\alpha) := \int_{D_{\phi_L}} e^{\alpha \phi_L(s)} ds = \int_{\phi^{-1}([0, L])} e^{\alpha \phi(s)} ds = \int_0^L e^{\alpha y} \psi(y) dy,$$

where the final equality above is a truncated analogue of (24), and follows from an application of (23). For  $t \in \mathbb{R}$  let  $\alpha_{t,L}$  denote the solution to  $\frac{\partial}{\partial \alpha} Z_L(\alpha)|_{\alpha=\alpha_{t,L}} = t$ .

Now by the triangle inequality, for any  $L > 0$  we have the inequalities

$$\|\mu_{\phi,t}^{N \rightarrow k} - \gamma_{\phi,\alpha_t}^{\otimes k}\| \leq \|\mu_{\phi_L,t}^{N \rightarrow k} - \gamma_{\phi_L,\alpha_{t,L}}^{\otimes k}\| + \|\mu_{\phi_L,t}^{N \rightarrow k} - \mu_{\phi,t}^{N \rightarrow k}\| + \|\gamma_{\phi_L,\alpha_{t,L}}^{\otimes k} - \gamma_{\phi,\alpha_t}^{\otimes k}\|$$

and

$$\|\mu_{\phi,t}^{N \rightarrow k} - \gamma_{\phi,\alpha_t}^{\otimes k}\| \leq \|\mu_{\phi_L,t}^{N \rightarrow k} - \gamma_{\phi_L,\alpha_{t,L}}^{\otimes k}\| - \|\mu_{\phi_L,t}^{N \rightarrow k} - \mu_{\phi,t}^{N \rightarrow k}\| - \|\gamma_{\phi_L,\alpha_{t,L}}^{\otimes k} - \gamma_{\phi,\alpha_t}^{\otimes k}\|.$$

In particular, we may write

$$\int_{\mathbb{R}^k} \left| \mu_{\phi,t}^{N \rightarrow k}(s) - \gamma_{\phi,\alpha_t}^{\otimes k}(s) \right| ds = \int_{\mathbb{R}^k} \left| \mu_{\phi_L,t}^{N \rightarrow k}(s) - \gamma_{\phi_L,\alpha_{t,L}}^{\otimes k}(s) \right| ds + \Delta_L, \quad (39)$$

where

$$|\Delta_L| \leq \int_{\mathbb{R}^k} |\gamma_{\phi_L,\alpha_{t,L}}^{\otimes k}(s) - \gamma_{\phi,\alpha_t}^{\otimes k}(s)| ds + \int_{\mathbb{R}^k} \left| \mu_{\phi_L,t}^{N \rightarrow k}(s) - \mu_{\phi,t}^{N \rightarrow k}(s) \right| ds. \quad (40)$$

We now show that for all  $\rho > 0$ , there exists an  $L_0$  such that whenever  $L \geq L_0$ ,  $|\Delta_L| < \rho$ . On the one hand, we note that the truncated Orlicz ball  $B_{\phi_L,t}^N$  is a subset of  $B_{\phi,t}^N$ , both of which have finite Lebesgue measure. Moreover, as  $L \rightarrow \infty$ ,  $B_{\phi,t}^N \setminus B_{\phi_L,t}^N \rightarrow \emptyset$ , so that by the continuity of Lebesgue measure we have  $|B_{\phi,t}^N \setminus B_{\phi_L,t}^N| \rightarrow 0$ . In particular, by Lemma 3.4 there is an  $L_1$  such that whenever  $L \geq L_1$ , we have

$$\int_{\mathbb{R}^k} \left| \mu_{\phi_L,t}^{N \rightarrow k}(s) - \mu_{\phi,t}^{N \rightarrow k}(s) \right| ds \leq \rho/2. \quad (41)$$

On the other hand, using Lemma 3.2 we have

$$\int_{\mathbb{R}^k} |\gamma_{\phi_L,\alpha_{t,L}}^{\otimes k}(s) - \gamma_{\phi,\alpha_t}^{\otimes k}(s)| ds \leq k \int_{\mathbb{R}} |\gamma_{\phi_L,\alpha_{t,L}}(s) - \gamma_{\phi,\alpha_t}(s)| ds \quad (42)$$

Now consider that  $\gamma_{\phi,\alpha_t}(s) = \frac{e^{\alpha_t \phi(s)}}{Z(\alpha_t)}$  and  $\gamma_{\phi_L,\alpha_{t,L}}(s) := \frac{e^{\alpha_{t,L} \phi(s)}}{Z_L(\alpha_{t,L})}$  are both densities that may be written as a function of  $\phi$ . In particular, we are in the setting of Lemma 3.5 with  $\Phi = \phi : \mathbb{R} \rightarrow [0, \infty]$ , so that

$$\int_{\mathbb{R}} |\gamma_{\phi_L,\alpha_{t,L}}(s) - \gamma_{\phi,\alpha_t}(s)| ds = \int_0^\infty |\phi^\# \gamma_{\phi_L,\alpha_{t,L}}(y) - \phi^\# \gamma_{\phi,\alpha_t}(y)| dy. \quad (43)$$

We now note that in the notation of Section 2, we have  $\phi^\# \gamma_{\phi_L, \alpha_{t,L}} = (\phi^\# \gamma_{\phi, \alpha_t})_{\langle L \rangle}$ , so that making the relevant substitution in (43) we have

$$\int_{\mathbb{R}} |\gamma_{\phi_L, \alpha_{t,L}}(s) - \gamma_{\phi, \alpha_t}(s)| ds = \int_0^\infty |(\phi^\# \gamma_{\phi, \alpha_t})_{\langle L \rangle}(y) - \phi^\# \gamma_{\phi, \alpha_t}(y)| dy.$$

Now by the second point in Lemma 2.2, there exists an  $L_2$  such that for every  $L \geq L_2$ ,

$$\int_0^\infty |(\phi^\# \gamma_{\phi, \alpha_t})_{\langle L \rangle}(y) - \phi^\# \gamma_{\phi, \alpha_t}(y)| dy \leq \rho/2k. \quad (44)$$

In particular, combining (44) and (42) with (41) in (40), we see that for all  $L \geq L_0 := \max\{L_1, L_2\}$ ,

$$|\Delta_L| \leq \rho. \quad (45)$$

Finally, since the truncated potential  $\phi_L$  has a domain of finite Lebesgue measure, by Lemma 6.3 we have

$$\int_{\mathbb{R}^k} \left| \mu_{\phi_L, t}^{N-k}(s) - \gamma_{\phi, \alpha_{t,L}}^{\otimes k}(s) \right| ds \begin{cases} \left(1 + \varepsilon_{k,N,L}^{(1)}\right) \xi_{k, \alpha_t} \frac{k}{N} & : k \text{ fixed}, N \rightarrow \infty, \\ \left(1 + \varepsilon_{k,N,L}^{(2)}\right) \xi \frac{k}{N} & : k = o(N), k, N \rightarrow \infty, \\ \left(1 + \varepsilon_{k,N,L}^{(3)}\right) Q(\theta) & : k \sim \theta N, k, N \rightarrow \infty, \end{cases} \quad (46)$$

where there is a universal constant  $C \in (0, \infty)$  such that if  $X_L$  is distributed according to  $(\gamma_{\alpha_t})_{\langle L \rangle}$ , then with  $C'_L := C \frac{\mathbb{E}[(X_L - \mathbb{E}[X_L])^4]}{\text{Var}[X_L]^2}$  we have

$$|\varepsilon_{k,N,L}^{(1)}| \leq C'_L \sqrt{\frac{k}{N}}, \quad |\varepsilon_{k,N,L}^{(2)}|, |\varepsilon_{k,N,L}^{(3)}| \leq C'_L \left( \frac{1}{\sqrt{k}} + \frac{1}{\sqrt{N-k}} \right).$$

Now by applying the third point of Lemma 2.2 to the truncated density  $(\gamma_{\alpha_t})_{\langle L \rangle}$ , we see that there exists a constant  $M \in (0, \infty)$  depending on  $\phi, t$  but independent of  $L$  such that for all  $L \geq L_0$  we have

$$\frac{\mathbb{E}[(X_L - \mathbb{E}[X_L])^4]}{\text{Var}[X_L]^2} \leq M,$$

so that for all  $L \geq L_0$  by setting  $C' = CM$  we have

$$|\varepsilon_{k,N,L}^{(1)}| \leq C' \sqrt{\frac{k}{N}}, \quad |\varepsilon_{k,N,L}^{(2)}|, |\varepsilon_{k,N,L}^{(3)}| \leq C' \left( \frac{1}{\sqrt{k}} + \frac{1}{\sqrt{N-k}} \right). \quad (47)$$

In particular, since  $\rho > 0$  and  $L$  are arbitrary, by combining (39), (45), (46) and (47) we obtain the result.  $\square$

## Appendix

In this appendix we sketch a few details of the proof of Lemma 4.2, which is very similar to the proof of Theorem 1.6 in [6].

*Proof of Lemma 4.2.* Since the proof runs almost identically to Section 3 of Diaconis and Freedman [6], we will only sketch the details in the case where  $k \sim \theta N$ .

Now, whenever  $\nu$  is a probability density on  $\mathbb{R}^k$  of a random variable  $(X_1, \dots, X_k)$ , let  $\bar{\nu}$  denote the probability density on  $\mathbb{R}^1$  of  $X_1 + \dots + X_k$ . According to the sufficiency Lemma 2.4 of [6], we have the equality in total variation

$$\|\psi_{\geq t}^{N-k} - \psi_{\alpha_t}^{\otimes k}\| = \|\bar{\psi}_{\geq t}^{N-k} - \bar{\psi}_{\alpha_t}^{\otimes k}\|.$$

Write  $f_k(s) := \bar{\psi}_{\alpha_t}^{\otimes k}(s)$ . We remark that if  $\psi^{\star k}(s)$  is the  $k$ -fold convolution of the density  $\psi$ , then

$$f_k(s) = \frac{e^{\alpha_t s} \psi^{\star k}(s)}{Z(\alpha_t)^k}. \quad (48)$$

Using (48) to obtain the second equality below, and the definition of  $\bar{\psi}_{\geq t}^{N-k}$  to obtain the first, the density associated with  $\bar{\psi}_{\geq t}^{N-k}$  is given by

$$\bar{\psi}_{\geq t}^{N-k}(s) = \frac{\psi^{\star k}(s) \int_{tN-s}^{\infty} \psi^{\star(N-k)}(u) du}{\int_{tN}^{\infty} \psi^{\star N}(u) du} = \frac{f_k(s) \int_{tN-s}^{\infty} e^{-\alpha(u-tN+s)} f_{N-k}(u) du}{\int_{tN}^{\infty} e^{-(u-tN)} f_N(u) du}.$$

In particular, we have

$$\|\bar{\psi}_{\geq t}^{N-K} - \bar{\psi}_{\geq t}^{\otimes k}\| = \int_{-\infty}^{\infty} f_k(s) \left| \frac{\int_{tN-s}^{\infty} e^{-\alpha(u-tN+s)} f_{N-k}(u) du}{\int_{tN}^{\infty} e^{-(u-tN)} f_N(u) du} - 1 \right| ds$$

We now consider a recentering. Let  $g_j$  denote the density of

$$\frac{X_1 + \dots + X_j - jt}{\sigma \sqrt{j}}$$

so that when  $j$  is large,  $g_j$  is close to the Gaussian density. Indeed, changing variable we have

$$\|\bar{\psi}_{\geq t}^{N-K} - \bar{\psi}_{\geq t}^{\otimes k}\| = \int_{-\infty}^{\infty} g_k(s) \left| \frac{\int_{-\phi s}^{\infty} e^{-\alpha \sigma \sqrt{N-k}(u+\phi s)} g_{N-k}(u) du}{\int_0^{\infty} e^{-\alpha \sigma \sqrt{N}u} g_N(u) du} - 1 \right| ds, \quad (49)$$

where  $\phi = \sqrt{\frac{k}{N-k}}$ . We now use the Edgeworth expansion to estimate the ratio

$$\frac{\int_{-\phi s}^{\infty} e^{-\alpha \sigma \sqrt{N-k}(u+\phi s)} g_{N-k}(u) du}{\int_0^{\infty} e^{-\alpha \sigma \sqrt{N}u} g_N(u) du} \quad (50)$$

as a function of the external variable  $s$ . Indeed, according to the Edgeworth expansion (see, e.g., [7, Section XVI])

$$g_j(s) = \frac{1}{\sqrt{2\pi}} e^{-s^2/2} \left( 1 + \frac{\kappa(\alpha_t)}{\sqrt{j}} (s^3 - 3s) \right) + \varepsilon_j(s), \quad (51)$$

where for a random variable  $X$  with density  $h_{\alpha_t}$ ,  $\kappa(\alpha_t) := \mathbb{E}[(X - \mathbb{E}[X])^3] / \text{Var}[X]^{3/2}$  and there is a universal constant  $C$  such that

$$|\varepsilon_j(s)| \leq \frac{\text{CE}[(X - \mathbb{E}[X])^4]}{\text{Var}[X]^2} \frac{1}{j}.$$

We begin with estimating the ratio (50) when  $k = \theta N$ . Here, it is easily seen that

$$\begin{aligned} \frac{\int_{-\phi s}^{\infty} e^{-\alpha \sigma \sqrt{N-k}(u+\phi s)} g_{N-k}(u) du}{\int_0^{\infty} e^{-\alpha \sigma \sqrt{N}u} g_N(u) du} &= \frac{\frac{1}{\alpha_t \sigma_t \sqrt{2\pi(N-k)}} \exp\left(-\frac{1}{2}(\phi s)^2\right) + O\left(\frac{1}{N-k}\right)}{\frac{1}{\alpha_t \sigma_t \sqrt{2\pi N}} + O(1/N)} \\ &= \frac{1}{\sqrt{1-\theta}} \exp\left(-\frac{1}{2} \frac{\theta}{1-\theta} s^2\right) + \varepsilon_{k,N}(s), \end{aligned} \quad (52)$$

where there is a universal constant  $C \in (0, \infty)$  such that for all  $s \in \mathbb{R}$  we have

$$|\varepsilon_{k,N}(s)| \leq C \frac{\mathbb{E}[|X - \mathbb{E}[X]|^4]}{\text{Var}[X]^{3/2}} \frac{1}{\sqrt{N-k}}$$

By plugging (52) into (49), and using the central limit theorem, we prove the statement of Lemma 4.2 in the setting where  $k \sim \theta N$ .

We omit the proof of the case  $k = o(N)$ , which is similar.  $\square$

## Acknowledgment

SJ and JP have been supported by the Austrian Science Fund (FWF) Project P32405 “Asymptotic Geometric Analysis and Applications” of which JP is principal investigator. JP is also supported by the Special Research Program Project F5508-N26.

## References

- [1] J. R. Blum and K. P. Pathak. A note on the zero-one law. *Ann. Math. Statist.*, 43:1008–1009, 1972.
- [2] E. Borel. *Introduction géométrique à quelques théories physiques*. Cornell University Library historical math monographs. Gauthier-Villars, 1914.
- [3] H. E. Daniels. Saddlepoint approximations in statistics. *Ann. Math. Statist.*, 25:631–650, 1954.
- [4] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications*. Stochastic Modelling and Applied Probability. Springer-Verlag Berlin Heidelberg, second edition, 2010.
- [5] P. Diaconis and D. Freedman. A dozen de Finetti-style results in search of a theory. *Ann. Inst. H. Poincaré Probab. Statist.*, 23(2, suppl.):397–423, 1987.
- [6] P. Diaconis and D. A. Freedman. Conditional limit theorems for exponential families and finite versions of de Finetti’s theorem. *J. Theoret. Probab.*, 1(4):381–410, 1988.
- [7] W. Feller. *An introduction to probability theory and its applications. Vol. II*. Second edition. John Wiley & Sons, Inc., New York-London-Sydney, 1971.
- [8] J. Hörrmann, J. Prochno, and C. Thäle. On the isotropic constant of random polytopes with vertices on an  $\ell_p$ -sphere. *J. Geom. Anal.*, 28(1):405–426, 2018.
- [9] Z. Kabluchko and J. Prochno. The maximum entropy principle and volumetric properties of Orlicz balls. *J. Math. Anal. Appl.*, 495(1):124687, 2021.
- [10] Z. Kabluchko, J. Prochno, and C. Thäle. High-dimensional limit theorems for random vectors in  $\ell_p^n$ -balls. *Commun. Contemp. Math.*, 21(1):1750092, 30, 2019.
- [11] Z. Kabluchko, J. Prochno, and V. Vysotsky. Yet another note on the arithmetic-geometric mean inequality. *Studia Math.*, 253(1):39–55, 2020.
- [12] S. S. Kim, Y.-T. Liao, and K. Ramanan. An asymptotic thin shell condition and large deviations for random multidimensional projections. *arXiv e-prints*, page arXiv:1912.13447, December 2019.
- [13] S. S. Kim and K. Ramanan. A conditional limit theorem for high-dimensional  $\ell^p$ -spheres. *J. Appl. Probab.*, 55(4):1060–1077, 2018.
- [14] A. Naor and D. Romik. Projecting the surface measure of the sphere of  $\ell_p^n$ . *Ann. Inst. H. Poincaré Probab. Statist.*, 39(2):241–261, 2003.
- [15] V. V. Petrov. On the probabilities of large deviations for sums of independent random variables. *Teor. Veroyatnost. i Primenen.*, 10:310–322, 1965.
- [16] J. Prochno, C. Thäle, and N. Turchi. The isotropic constant of random polytopes with vertices on convex surfaces. *J. Complexity*, 54:101394, 17, 2019.

- [17] S. T. Rachev and L. Rüschendorf. Approximate independence of distributions on spheres and their stability properties. *Ann. Probab.*, 19(3):1311–1337, 1991.
- [18] G. Schechtman and J. Zinn. On the volume of the intersection of two  $L_p^n$  balls. *Proc. Amer. Math. Soc.*, 110(1):217–224, 1990.

SAMUEL G. G. JOHNSTON: Institute of Mathematics and Scientific Computing, University of Graz, Heinrichstrasse 36, 8010 Graz, Austria  
*E-mail:* samuel.johnston@uni-graz.at

JOSCHA PROCHNO: Institute of Mathematics and Scientific Computing, University of Graz, Heinrichstrasse 36, 8010 Graz, Austria  
*E-mail:* joscha.prochno@uni-graz.at