A Maxwell principle for generalized Orlicz balls

Samuel G. G. Johnston\textsuperscript{a} and Joscha Prochno\textsuperscript{b}

\textsuperscript{a}Department of Mathematical Sciences, University of Bath, Claverton Down, Bath BA2 7AY, United Kingdom. E-mail: sgj22@bath.ac.uk
\textsuperscript{b}Faculty of Computer Science and Mathematics, University of Passau, Innstraße 33, 94032 Passau, Germany. E-mail: joscha.prochno@uni-passau.de

\textbf{Abstract.} In the 1980s, 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 vectors respectively. In subsequent works, Racah 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 similar and significantly broader setting, looking at low-dimensional projections of random vectors uniformly distributed on sets of the form $B^n_N := \{(s_1, \ldots, s_N) \in \mathbb{R}^N : \sum_{i=1}^N s_i^2 \leq t N\}$, where $\phi : \mathbb{R} \rightarrow [0, \infty]$ is a function satisfying some fairly mild conditions. 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 $B^n_N$ behave like independent 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 $e^{-\beta_t \phi(s)}$.

\textbf{MSC2020 subject classifications:} Primary 60F05; secondary 52A20, 60F10

\textbf{Keywords:} generalized Orlicz balls, Gibbs conditioning principle, Gibbs measures, Maxwell principle, low-dimensional projections, quantitative Cramér theorem

1. Introduction

1.1. The Maxwell principle

Over a century ago, Borel [4, Chapter 5] observed 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. Borel attributed this result to Maxwell, and as such this result is commonly known as the Maxwell-Borel lemma or the Maxwell principle. More precisely, consider the Euclidean sphere

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

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 $(\Theta_1, \ldots, \Theta_N)$ is a random vector chosen according to $\sigma^N$, the unique rotationally invariant probability measure on $S_2^{N-1}$. For $k < N$, let $\sigma^{N-k}$ denote the probability density function on $\mathbb{R}^k$ associated with the marginal law of the first $k$ coordinates $(\Theta_1, \ldots, \Theta_k)$ of $(\Theta_1, \ldots, \Theta_N)$, which is given by

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

where $\mathbb{1}\{s_1^2 + \ldots + s_k^2 \leq N\}$ is the indicator function of the event $\{s_1^2 + \ldots + s_k^2 \leq N\}$. It turns out that when $N$ is large and $k$ is small compared to $N$, the probability density $\sigma^{N-k}$ on $\mathbb{R}^k$ is very close to the $k$-dimensional product $\gamma^{\otimes k}$ of the standard Gaussian density

$$\gamma(s) := \frac{1}{\sqrt{2\pi}} e^{-s^2/2}, \quad s \in \mathbb{R}.$$
(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 [7] 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 related phenomena in different settings. Indeed, in place of $S_{2}^{N-1}$ take instead the simplex

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

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 Rüschendorf [20] connected these observations in the setting of the $\ell_p^N$-spheres

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

establishing an $\ell_p$-version of the classical Maxwell-Borel lemma involving the standard $p$-Gaussian density

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

Rachev and Rüschendorf 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 $\gamma_p^N$, then the normalised vector $\zeta/||\zeta||_p$ takes values in the $p$-sphere and is distributed according to $\mu_p^N$; see also [21, Lemma 1] or [18, Proof of Lemma 4]. Rachev and Rüschendorf use this probabilistic construction to analyze the marginal density $\mu^{N\rightarrow k}_p$ of the first $k$ coordinates $(X_1, \ldots, X_k)$ of a random vector $(X_1, \ldots, X_N)$ distributed according to the cone measure $\mu_p^N$. Indeed, they show that when $N \to \infty$ with $k = o(N)$, we have the total variation estimate

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

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 [20] beautifully synthesises 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 [18] 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 obtain the total variation bound

$$\int_{\mathbb{R}^k} |\sigma^{N\rightarrow k}_p(s) - \gamma^{\otimes k}_p(s)| \, ds \leq C_p \left( \frac{k}{N} + \frac{1}{\sqrt{N}} \right)$$

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 [20] and Naor and Romik [18] certainly provide a gratifying answer to Diaconis and Freedman’s [7] appeal for a unified theory, the purpose of the present paper is to show that these phenomena may be unified in a far broader framework in the setting of generalized Orlicz balls

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

where \( t > 0 \) and \( \phi : \mathbb{R} \rightarrow [0, \infty) \) is a potential — a measurable 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, 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 however the sets \( B_{\phi,t}^N \) we consider need not be convex, symmetric, simply connected or even compact. This general setting is made possible through a different perspective based on 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 the 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 \( D_{\phi} \).

- 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 the partition function, and \( \beta_t > 0 \) is a parameter (in statistical mechanics parlance referred to as inverse temperature) chosen so that

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

This tells us that the coordinates of the ball \( B_{\phi,t}^N \) 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

Shortly we will 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. For measurable functions \( \phi : \mathbb{R} \rightarrow [0, \infty] \), we define the domain \( D_{\phi} := \{ s \in \mathbb{R} : \phi(s) < \infty \} \) to be the set of points where it is finite.
Definition 1.1. A measurable function $\phi : \mathbb{R} \rightarrow [0, \infty]$ is a potential if it is differentiable on its domain $D_\phi$ with the possible exception of a countable subset of $D_\phi$, the essential infimum of $\phi$ is zero, the level sets $\{ s \in \mathbb{R} : \phi(s) = y \}$ of $\phi$ are finite for $y < \infty$, and the partition function $Z : \mathbb{R} \rightarrow (-\infty, +\infty]$ of $\phi$, given by

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

satisfies the technical condition that $(-\infty, \alpha_{\sup}) := \{ \alpha : Z(\alpha) < \infty \}$ is a non-empty open interval, and that moreover we have $\lim_{\alpha \uparrow \alpha_{\sup}} Z(\alpha) = \infty$.

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 $D_\phi \phi, y$ 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 $B_{\phi,t}^\infty$.

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 $B_p^N$ ball of radius $(tN)^{1/p}$, and the function $\phi(s) := \infty 1_{(-\infty, 0)}(s) + s$ corresponds to the simplex $D_\phi^{N-1}$ 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 [12, Equation (3)].

We note in particular that if $Z(\alpha) < \infty$ for all negative $\alpha$, but $Z(0) = \infty$ (i.e. $D_\phi$ has infinite Lebesgue measure), then by the monotone convergence theorem the technical condition is satisfied. In particular, if for some $\lambda, \varepsilon > 0$, $\phi$ satisfies $\lambda |s|^\varepsilon \leq \phi(s) < \infty$ for all real $s$, then the technical condition is satisfied.

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 set with none of these properties, we invite the reader to consider the set $B_{\phi,t}^N$ associated with the potential

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

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.

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

It is easily verified using the fact that $Z(\alpha) < \infty$ for some $\alpha$ that $|D_\phi| = \infty$ implies $t_{\sup} = \infty$. Our second quantity, which we already mentioned above, is the domain average of $\phi$, given by

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

We emphasise that the integral $\int_{D_\phi} \phi(s) \, ds$ may be infinite even when $|D_\phi| < \infty$, and that $t_{\crit} \leq t_{\sup}$ in any case.

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

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

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

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

for the uniform density on the domain of $\phi$. 
Let us remark that
\[ \frac{\partial}{\partial \alpha} \log Z(\alpha) = \frac{\int_{D_\phi} \phi(s) e^{\alpha \phi(s)} ds}{\int_{D_\phi} e^{\alpha \phi(s)} ds} = \int_{-\infty}^\infty \phi(s) \gamma_{\phi, \alpha}(s) ds. \]

The following lemma guarantees for every \( t \in (0, t_{sup}) \) the existence of a parameter \( \alpha_t \) such that the expectation of \( \phi \) against \( \gamma_{\phi, \alpha_t}(s) \) is equal to \( t \).

**Lemma 1.3.** If \( \phi \) is a potential then for each \( t \in (0, t_{sup}) \) there exists a unique parameter \( \alpha_t \in \mathbb{R} \) such that
\[ W(\alpha_t) := \frac{Z'(\alpha_t)}{Z(\alpha_t)} = \int_{-\infty}^\infty \phi(s) \gamma_{\phi, \alpha_t}(s) ds = t. \]
Moreover, if \( t < t_{crit} \), then \( \alpha_t < 0 \), whereas if \( t > t_{crit} \), \( \alpha_t > 0 \).

Results of this form are well known (for Orlicz functions we refer to [10, Lemma 3.1] as well as [1, 12]), though since we could not find a reference for our exact formulation, we provide a short proof in the Appendix, adapting the proof of [5, Theorem 6.2].

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, \( \mu^{N \to k}_{\phi, t} \) denotes the marginal density of the first \( k \) coordinates \( (X_1, \ldots, X_k) \) of a random vector \( (X_1, \ldots, X_N) \) chosen according to the uniform measure \( \mu^N_{\phi, t} \) on the generalized Orlicz ball \( B^N_{\phi, t} \). That is
\[ \mu^{N \to k}_{\phi, t}(s_1, \ldots, s_k) := \frac{1}{|B^N_{\phi, t}|} \int_{\mathbb{R}^{N-k}} 1_{B^N_{\phi, t}}(s_1, \ldots, s_k, s_{k+1}, \ldots, s_N) ds_{k+1} \ldots ds_N. \]

With \( t_{crit} \) and \( t_{sup} \) now defined, shortly we state our main result, Theorem 1.4. Beforehand however, let us briefly highlight that the case \( t \geq t_{sup} \) is trivial. Here, the generalized Orlicz ball is identical to the \( N \)-fold product of the domain of \( \phi \), that is \( B^N_{\phi, t} = D^N_{\phi} = \{(s_1, \ldots, s_N) \in \mathbb{R}^N : s_i \in D_{\phi}\} \), and hence the coordinates of \( B^N_{\phi, t} \) are independent and uniformly distributed on \( D_{\phi} \). In particular, \( \mu^{N \to k}_{\phi, t} = \gamma_{\phi, uni}^{\otimes k} \).

Our main statement is concerned with the low-dimensional projections of \( B^N_{\phi, t} \) in the setting \( t < t_{sup} \). Here we find that a phase transition occurs at the point \( t = t_{crit} \).

**Theorem 1.4.** If \( t_{crit} < t < t_{sup} \), then the \( k \)-dimensional projections of a random vector uniformly sampled from \( B^N_{\phi, t} \) are close in distribution to a \( k \)-dimensional product of \( \gamma_{\phi, uni} \). 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
\[ \left| \int_{\mathbb{R}^k} \left| \mu^{N \to k}_{\phi, t}(s) - \gamma_{\phi, uni}^{\otimes k}(s) \right| ds \right| \leq C e^{-cN}. \]

On the other hand, if \( t \leq t_{crit} \), then the \( k \)-dimensional projections of a random vector uniformly sampled from \( B^N_{\phi, t} \) 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
\[ \left| \int_{\mathbb{R}^k} \left| \mu^{N \to k}_{\phi, t}(s) - \gamma_{\phi, \alpha_t}^{\otimes k}(s) \right| ds \right| \leq C \frac{k}{N}. \]

In fact, Theorem 1.4 has been abbreviated for the sake of clarity, and is a condensed version of finer statements which we state and prove in the sequel. More specifically, in Sections 5 and 6 we study separately the respective cases \( t_{crit} < t \) and \( t \leq t_{crit} \), and in both settings we obtain fine estimates of the total variations occurring on the left-hand sides of (5) and (6).

### 1.4. Further discussion

Our approach to proving Theorem 1.4 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 [18] and [20] 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 [8]. Somewhat surprisingly this result was neither cited in [18] nor [20] (or the recent paper [16]) even though it already contains ideas towards a unified and generalized theory in the sense of Rachev and Rüschendorf [20]. But this exact paper [8] 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., [12] and [15, 16].

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 1.4. 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

\[
\frac{\left| B_{\phi,t}^N \right|}{\left| D_{\phi}^N \right|} \to 1, \quad \text{as} \quad N \to \infty
\]

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 Cramér’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 \( t < t_{\text{crit}} \). Here we find that in contrast to (7), either \( \left| D_{\phi} \right| = \infty \), or even when \( \left| D_{\phi} \right| < \infty \), we have

\[
\frac{\left| B_{\phi,t}^N \right|}{\left| D_{\phi}^N \right|} \to 0, \quad \text{as} \quad N \to \infty.
\]

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 \( Y_1, Y_2, \ldots \) are independent and identically distributed random variables distributed according to a probability density \( \nu \) on the real line, and suppose further that \( \mathbb{E}[Y_1] = t_0 \). For \( t < t_0 \), the Gibbs conditioning principle is concerned with the asymptotic distribution of first \( k \) coordinates \( (Y_1, \ldots, Y_k) \) conditioned on the large deviation event \( \{ Y_1 + \ldots + Y_N \leq tN \} \). 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

\[
\nu_{\alpha_k}(y) := e^{\alpha_k y} \nu(y)/Z(\alpha_k), \quad s \in \mathbb{R},
\]

where \( Z(\alpha) \) is the moment generating function associated with \( \nu \) and \( \alpha < 0 \) is a parameter chosen so that \( \int_{-\infty}^{\infty} y \nu_{\alpha_k}(y) dy \) is equal to \( t \).

With a view to relating the Gibbs conditioning principle to our problem, in the setting where \( \left| D_{\phi} \right| \) is finite, we take a sequence of random variables \( X_1, X_2, \ldots \) sampled uniformly and independently from \( D_{\phi} \), and consider the transformed sequence \( Y_1, Y_2, \ldots \) given by \( Y_i := \phi(X_i) \). This transformation allows us to express membership of a random vector \( (X_1, \ldots, X_N) \) in the subset \( B_{\phi,t}^N \) of \( D_{\phi}^N \) in terms of the large deviation event \( \{ \phi(X_1) + \ldots + \phi(X_N) \leq Nt \} = \{ Y_1 + \ldots + Y_N \leq Nt \} \), and therefore use the Gibbs conditioning principle to understand the low-dimensional projections of \( (Y_1, \ldots, Y_k) \), and hence ultimately \( (X_1, \ldots, X_k) \) on this event. 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 \( \alpha_k = 0 \). 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 1.4 with the function \( \phi(s) = s^p \) and the frameworks considered by Diaconis and Freedman [7], Rachev and Rüschendorf [20], and Naor and Romik [18]. Namely, the above authors consider sampling from the boundary of the \( \ell_p^N \) ball (with either the cone or surface measure), while for aesthetic reasons we consider sampling uniformly from the interior of our generalized Orlicz balls \( B_{\phi,t}^N \). Ultimately we recover an analogous result in our slightly different setting, namely that the low-dimensional projections of \( \ell_p^N \) balls are approximately \( p \)-Gaussian. That said, we conjecture that under some mild conditions, if we were to consider sampling from either the surface or cone measure on the boundary of the generalized Orlicz ball \( B_{\phi,t}^N \) (as opposed to the sampling from the uniform distribution on its interior), we would continue to see the phenomena that in high dimensions the low dimensional projections behave like independent random variables with laws
Suppose \( \gamma_{\alpha,L}(s) \), in analogy with Theorem 1.4. We do not provide evidence for this conjecture here, but are content to remark 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., [11, 13, 14, 17, 19, 21]). 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 Cramér’s theorem and the Gibbs conditioning principle that are used in the following two sections to prove Theorem 1.4.
- In Section 5 we give a proof of the case \( t > t_{\text{crit}} \) of Theorem 1.4.
- In Section 6 we give a proof of the case \( t \leq t_{\text{crit}} \) of Theorem 1.4.
- In the Appendix of the paper we give a proof of the quantitative version of the Gibbs conditioning principle, Lemma 4.4.

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 \( \nu : [0, \infty) \to [0, \infty) \) is a function such that there exist constants \( c, C \in (0, \infty) \) such that

\[
\nu(y) \leq Ce^{-cy}.
\]

Then, for all \( \kappa \geq 0 \), and all \( L > \kappa/c \), we have

\[
\int_L^\infty y^\kappa \nu(y) \, dy \leq \frac{C L^\kappa e^{-cL}}{c - \kappa/L} \leq C' e^{-c'L}
\]

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

**Proof.** We have

\[
\int_L^\infty y^\kappa \nu(y) \, dy \leq C \int_L^\infty y^\kappa e^{-cy} \, dy = C L^\kappa e^{-cL} \int_0^\infty \left( 1 + \frac{y}{L} \right)^\kappa e^{-cy} \, dy.
\]

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

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 \( \nu_{(L)} \) of the measure restricted to taking values in \( [0, y) \), but exponentially tilted so that \( \nu_{(L)} \) has the same mean as \( \nu \). The following lemma states that the measure \( \nu_{(L)} \) 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 \( \nu : [0, \infty) \to [0, \infty) \) be a probability density satisfying \( \nu(y) \leq Ce^{-cy} \) for all \( y \geq 0 \), with expectation \( \int_0^\infty y\nu(y) \, dy = t \) and moment generating function \( Z(\alpha) := \int_0^\infty e^{\alpha y}\nu(y) \, dy \) (which is clearly finite for at least all \( \alpha < c \)).

Given \( L \in (0, \infty) \) and \( \alpha \in \mathbb{R} \), we define the tilted truncation \( \nu_{\alpha,L} \) to be the density function

\[
\nu_{\alpha,L}(y) := \frac{e^{\alpha y}1_{\{y < L\}}\nu(y)}{\int_0^L e^{\alpha y}\nu(y) \, dy}.
\]

Then there exists \( L_0 \in (0, \infty) \) such that for all \( L \geq L_0 \), there is a unique parameter \( \alpha(L) \geq 0 \) such that \( \nu_{\alpha(L),L} \) satisfies \( \int_0^L y\nu_{\alpha(L),L}(y) \, dy = t \). Moreover, if we write \( \nu_{(L)} \) for \( \nu_{\alpha(L),L} \), there are constants \( C', c' \in (0, \infty) \) such that we have the bounds:
1. The size of \( \alpha(L) \) is bounded by \( \alpha(L) \leq C'e^{-c'L} \).

2. The total variation between \( \nu(L) \) and \( \nu \) is bounded by

\[
\|\nu(L) - \nu\| := \int_0^\infty |\nu(L)(y) - \nu(y)| \, dy \leq C'e^{-c'L}.
\]

3. The difference between the \( j \)-th moment of \( \nu \) and \( \nu(L) \) is bounded by

\[
\left|\int_0^\infty y^j \nu(y) \, dy - \int_0^\infty y^j \nu(L)(y) \, dy\right| \leq \frac{C'}{1 - j/L_0} L^{j+1} e^{-c'L}.
\]

**Proof.** In the case that \( \nu \) has compact support in \([0, L_0]\), say, then for all \( L \geq L_0 \) we can take \( \alpha(L) = 0 \) and \( \nu(L) \) 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 compactly supported.

We begin by establishing the existence and uniqueness of the parameter \( \alpha(L) \). Define \( Z_L(\alpha) := \int_0^L e^{\alpha y} \nu(y) \, dy \), so that \((\log Z_L)'(\alpha) = \int_0^L y e^{\alpha y} \nu(y) \, dy \). Moreover, for \( t \in \mathbb{R} \) we set \( Z_L(\alpha, t) := Z_L(\alpha) e^{-\alpha t} := \int_0^L e^{\alpha(y-t)} \nu(y) \, dy \), then it is easily verified that \( Z_L'(\alpha, t) = 0 \) implies \((\log Z_L)'(\alpha) = \int_0^L y \nu(y) \, dy = t \). Consequently, it remains to establish the existence and uniqueness of a solution \( \alpha \) to the equation \( Z_L'(\alpha, t) = 0 \).

Now since \( Z_L'(\alpha, t) = \int_0^\infty (y-t)^2 e^{\alpha(y-t)} \, dy > 0 \), \( Z_L'(\alpha, t) \) is strictly increasing in \( \alpha \). Now with \( t := \int_0^\infty y \nu(y) \, dy \), there exist \( \varepsilon, \varepsilon' > 0 \) such that \( \int_{-\varepsilon}^{\varepsilon} \nu(y) \, dy > \varepsilon' \). Let \( L_0 = t + \varepsilon \). We now establish that for all \( L \geq L_0 \), we have

\[
\lim_{\alpha \downarrow -\infty} Z_L'(\alpha, t) = -\infty \quad \text{and} \quad \lim_{\alpha \uparrow \infty} Z_L'(\alpha, t) = \infty.
\]

Note that \( Z_L'(\alpha, t) = \int_0^\infty (y-t) e^{\alpha(y-t)} \nu(y) \, dy \). Now on the one hand, since \( \int_0^\infty y \nu(y) \, dy = t \), it follows that \( \int_0^L \nu(y) \, dy \geq 0 \). By the monotone convergence theorem it follows that \( \lim_{\alpha \downarrow -\infty} Z_L'(\alpha, t) = \lim_{\alpha \downarrow -\infty} \int_0^\infty (y-u) e^{\alpha(y-u)} \nu(y) \, dy = -\infty \). On the other hand, since \( \int_0^L \nu(y) > 0 \), again by the monotone convergence theorem we have \( \lim_{\alpha \uparrow \infty} Z_L'(\alpha, t) = \lim_{\alpha \uparrow \infty} \int_0^L (y-u) e^{\alpha(y-u)} \nu(y) \, dy = \infty \). That proves (9). We have now seen that for \( L \geq L_0 \), as a function of \( \alpha \), \( Z_L'(\alpha, t) \) is a strictly increasing bijection from \( \mathbb{R} \) to \( \mathbb{R} \). In particular there exists a unique \( \alpha(L) \) such that \( Z_L'(\alpha(L), t) = 0 \), and hence \((\log Z_L)'(\alpha(L)) = t \).

We now prove that \( \alpha(L) \geq 0 \). To this end, since \( Z_L'(\alpha, t) \) is increasing in \( \alpha \), it is sufficient to establish that \( Z_L'(0, t) \leq 0 \).

We now note using the definition \( t := \int_0^\infty y \nu(y) \, dy \) to obtain the second equality below, and then the fact that \( L \geq L_0 > t \) to obtain the following inequality, we have

\[
Z_L'(0, t) = \int_0^L (y-t) \nu(y) \, dy = -\int_L^\infty (y-t) \nu(y) \, dy \leq 0.
\]

Next, we show that when \( L \) is large, \( \alpha(L) \) is small. First we note that the generating function \( Z(\alpha) := \int_0^\infty e^{\alpha y} \nu(y) \, dy \) exists in a subset of \( \mathbb{R} \) containing \((-\infty, c)\) and satisfies \( Z(0) = 1 \), \( Z'(0) = t \) and \( Z''(0) > t^2 \) (the latter holding since \( \nu \) is non-degenerate). Now since the truncated moment generating function \( Z_L(\alpha) := \int_0^L e^{\alpha y} \nu(y) \, dy \) is logarithmically convex, the derivative of its logarithm \( Q_L(\alpha) := (\log Z_L)'(\alpha) \) is increasing in the \( \alpha \) variable. By our discussion above, \( Q_L(\alpha(L)) = t \).

Now \( Q_L(\alpha) \) takes the form

\[
Q_L(\alpha) := \frac{\int_0^L y e^{\alpha y} \nu(y) \, dy}{\int_0^L e^{\alpha y} \nu(y) \, dy}.
\]

Next, we develop a lower bound for \( Q_L(\alpha) \). Extending the range of the integral in the denominator to obtain the inequality below, we have

\[
Q_L(\alpha) = \frac{\int_0^\infty y e^{\alpha y} \nu(y) \, dy - \int_L^\infty y e^{\alpha y} \nu(y) \, dy}{\int_0^L e^{\alpha y} \nu(y) \, dy} \\
\geq \frac{\int_0^\infty y e^{\alpha y} \nu(y) \, dy - \int_L^\infty y e^{\alpha y} \nu(y) \, dy}{\int_0^\infty e^{\alpha y} \nu(y) \, dy}.
\]
\[ Q(\alpha) = \int_{L}^{\infty} ye^{\alpha y} \nu(y) \, dy - \int_{0}^{\infty} e^{\alpha y} \nu(y) \, dy, \quad \text{(11)} \]

where \( Q(\alpha) := \frac{\partial}{\partial \alpha} \log Z(\alpha) \). With a view to bounding \( Q_L(\alpha) \) below, we now control the two quantities appearing on the right-hand-side of (11).

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

\[ Q(\alpha) \geq t + \rho \alpha. \quad \text{(12)} \]

As for the integral in the final line of (11), recall that \( \nu(y) \leq Ce^{-cy} \) for all \( y \). In particular, setting \( f(y) = e^{\alpha y} \nu(y) \) in Lemma 2.1, we see that there are constants \( c, C \) such that for all \( \alpha \in [0, c/2] \),

\[ \int_{L}^{\infty} ye^{\alpha y} \nu(y) \, dy \leq C_1 e^{-c_1 L}. \quad \text{(13)} \]

Combining (12) and (13) in (11), we obtain

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

Now increase \( L_0 \) if necessary, to ensure that \( L_0 \) is 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

\[ Q_L(\alpha) \geq t + \rho \alpha - C_1 e^{-c_1 L} \quad \text{for } \alpha \in \left[0, \frac{C_1}{\rho} e^{-c_1 L}\right]. \quad \text{(14)} \]

It follows that for all \( L \geq L_0 \), \( Q_L \left( \frac{C_1}{\rho} e^{-c_1 L} \right) \geq t \), 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 \( \nu(L) \) is supported on \([0, L]\), we have

\[ \int_{0}^{\infty} |\nu(L)(y) - \nu(y)| \, dy = \int_{0}^{L} \left| \frac{e^{\alpha(L)y}}{\int_{0}^{L} e^{\alpha(L)y} \nu(y) \, dy} - 1 \right| \nu(y) \, dy + \int_{L}^{\infty} \nu(y) \, dy. \quad \text{(15)} \]

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 (15) may be bounded by

\[ \int_{L}^{\infty} \nu(y) \, dy \leq C_2 e^{-c_2 L}. \quad \text{(16)} \]

We turn to bounding the former integral on the right hand side of (15). 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

\[ \begin{align*}
1 & \leq e^{\alpha(L)y} \leq e^{\frac{C_1}{\rho} Le^{-c_1 L}} \leq 1 + C_3 e^{-c_3 L}.
\end{align*} \quad \text{(17)} \]

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 (17) to obtain the upper bound below, it may be seen that there are constants \( C_4, c_4 \in (0, \infty) \) such that

\[ \begin{align*}
1 - C_4 e^{-c_4 L} & \leq \int_{0}^{L} e^{\alpha(L)y} \nu(y) \, dy \leq 1 + C_4 e^{-c_4 L}.
\end{align*} \quad \text{(18)} \]

Combining (17) with (18), we see that there exist constants \( C_5, c_5 \in (0, L) \) such that for every \( s \in [0, L] \),

\[ \left| \frac{e^{\alpha(L)y}}{\int_{s}^{L} e^{\alpha(L)y} \nu(y) \, dy} - 1 \right| \leq C_5 e^{-c_5 L}. \quad \text{(19)} \]
Using the fact that \( \nu \) is a probability measure, (19) entails
\[
\int_0^L \left| \frac{e^{\alpha(L)y}}{\int_0^L e^{\alpha(L)y} \nu(y) dy} - 1 \right| \nu(y) dy \leq C_7 e^{-c_5 L}.
\] (20)

In particular, using the bounds (16) and (20) in (15), 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 \( \nu(L) \), the proof of which follows quickly from the bound (19). Indeed, using the triangle inequality to obtain the first inequality below, and (19) to obtain the second we have
\[
\left| \int_0^\infty y^j \nu(y) dy - \int_0^\infty y^j \nu(L)(y) dy \right| \leq C_5 e^{-c_5 L} \int_0^L y^j \nu(y) dy + \int_0^\infty y^j \nu(y) dy,
\]
\[
\leq C_5 e^{-c_5 L} \int_0^L y^j \nu(y) dy + C_6 e^{-c_5 L},
\]
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. It particular there exist \( C_7, c_7 \in (0, \infty) \) such that
\[
\left| \int_0^\infty y^j \nu(y) dy - \int_0^\infty y^j \nu(L)(y) dy \right| \leq C_7 e^{-c_7 L},
\]
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 \( \pi^{N-k} \) for the marginal density on \( \mathbb{R}^k \) of the first \( k \) coordinates \( (X_1, \ldots, X_k) \), where \( (X_1, \ldots, X_N) \) 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
\[
||\pi - \lambda|| = \int_E |f - g| d\nu
\] (21)
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 [3].

**Lemma 3.2** (Total variation summing lemma). *Let \( \pi \) and \( \lambda \) be probability densities on \( \mathbb{R} \), and let \( \pi^{\otimes k} \) and \( \lambda^{\otimes k} \) be their respective product densities on \( \mathbb{R}^k \). Then*
\[
||\pi^{\otimes k} - \lambda^{\otimes k}|| \leq k||\pi - \lambda||.
\]
Proof. We prove the claim by induction. The result is trivial for $k = 1$. For $k \geq 2$, using the triangle inequality to obtain the first inequality below, and the fact that $\pi$ and $\lambda^{\otimes (k-1)}$ are probability densities to obtain the final equality, we have

$$
\|\pi^{\otimes k} - \lambda^{\otimes k}\| = \int_{\mathbb{R}^k} |\pi(s_1) \ldots \pi(s_k) - \lambda(s_1) \ldots \lambda(s_k)| ds_1 \ldots ds_k
\leq \int_{\mathbb{R}^k} |\pi(s_1) \ldots \pi(s_k) - \lambda(s_1) \ldots \lambda(s_{k-1})\pi(s_k)| ds_1 \ldots ds_k
+ \int_{\mathbb{R}^k} |\lambda(s_1) \ldots \lambda(s_{k-1})\pi(s_k) - \lambda(s_1) \ldots \lambda(s_k)| ds_1 \ldots ds_k
= \int_{\mathbb{R}^k} |\pi(s_1) \ldots \pi(s_{k-1}) - \lambda(s_1) \ldots \lambda(s_{k-1})|\pi(s_k)| ds_1 \ldots ds_k
+ \int_{\mathbb{R}^k} \lambda(s_1) \ldots \lambda(s_{k-1})|\pi(s_k)| ds_1 \ldots ds_k
= ||\pi^{\otimes (k-1)} - \lambda^{\otimes (k-1)}|| + ||\pi - \lambda||.
$$

The result now follows from the inductive hypothesis.

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 $\pi^{N \to k}$ and $\lambda^{N \to k}$ as in Definition 3.1 we have

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

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

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

which completes the proof.

Our next lemma states that if $A \subseteq B$, and $B \setminus A$ has small Lebesgue measure, 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 $\mu_A^{N \to k}$ and $\mu_B^{N \to k}$ defined in Definition 3.1, we have

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

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

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

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 \(\Phi : E \to F\) 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 \(F\), where \(\Phi^{-1} (B) := \{ e \in E : \Phi (e) \in B \}\). 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, [7, Lemma (2.4)], stating that 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 \(\Phi : E \to F\) 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 \(f, g : F \to [0, \infty]\) such that

\[
\left| \frac{d \pi}{d \mu} \right| = f \circ \Phi \quad \text{and} \quad \left| \frac{d \lambda}{d \mu} \right| = g \circ \Phi.
\]

Then

\[
\| \pi - \lambda \| = \| \Phi^\# \pi - \Phi^\# \lambda \|.
\] (22)

**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 (21) to obtain the outer equalities below, and changing variable to obtain the central equality, we have

\[
\| \Phi^\# \pi - \Phi^\# \lambda \| = \int_{E} |f(\Phi(y)) - g(\Phi(y))| \Phi^\# \mu(\mathrm{d}y) = \int_{E} |f(\Phi(s)) - g(\Phi(s))| \mu(\mathrm{d}s) = \| \pi - \lambda \|
\]

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 \(f : \mathbb{R} \to \mathbb{R}\) is a measurable mapping, we write \(f^\# \nu\) for the probability density on \(\mathbb{R}\) associated with the pushforward by \(f\) of the measure \(\nu(\mathrm{d}s)\). 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 \(F : [0, \infty) \to [0, \infty)\) 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

\[
\psi(y) := F'(y) = \sum_{s \in \phi^{-1}(y)} 1/|\phi'(s)|,
\] (23)

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,

\[
\int_{-\infty}^{\infty} f(\phi(s)) \mathrm{d}s = \int_{0}^{\infty} f(y) \psi(y) \mathrm{d}y.
\] (24)

In particular, whenever \(\nu : \mathbb{R} \to [0, \infty)\) is a probability density of the form \(\nu(s) = f(\phi(s))\), the pushforward \(\phi^\# \nu\) of the measure \(\nu(\mathrm{d}s)\) 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 (24) we may alternatively write

$$Z(\alpha) := \int_0^\infty e^{\alpha y} \psi(y) dy. \quad (25)$$

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

$$\psi_\alpha(y) := \frac{e^{\alpha y} \psi(y)}{Z(\alpha)}. \quad (26)$$

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

$$\psi_0(y) := \frac{\psi(y)}{|D_\phi|}. \quad (27)$$

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

Finally, define the multivariate potential $\Phi : \mathbb{R}^k \to [0, \infty)^k$ by $\Phi(s_1, \ldots, s_k) := (\phi(s_1), \ldots, \phi(s_k))$. We note that whenever $\nu$ is a measure on $\mathbb{R}$, we have

$$(\phi^# \nu)^\otimes k = \Phi^# (\nu^\otimes k). \quad (28)$$

In other words, if $X$ is a random variable with law $\nu$, then $\phi(X)$ has law $\phi^# \nu$. It follows that if $(X_1, \ldots, X_k)$ is a vector of random variables identically distributed like $\nu$, both sides of (28) express the law of the random vector $(\phi(X_1), \ldots, \phi(X_k)) = \Phi(X_1, \ldots, X_k)$.

### 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 1.4. Throughout this section, we will work under the following assumption.

**Assumption 4.1.** Let $Y_1, Y_2, \ldots$ be a sequence of independent random variables identically distributed according to a probability density $\psi : \mathbb{R} \to [0, \infty)$ whose generating function

$$Z(\alpha) := \int_{-\infty}^\infty e^{\alpha y} \psi(y) dy$$

is defined on a non-empty interval $I$ of $\mathbb{R}$ containing an open neighbourhood of zero.

We furthermore assume that if $\alpha_{\inf} := \inf I$ and $\alpha_{\sup} = \sup I$, the technical condition

$$\lim_{\alpha \downarrow \alpha_{\inf}} Z(\alpha) = -\infty \quad \text{and} \quad \lim_{\alpha \uparrow \alpha_{\sup}} Z(\alpha) = +\infty$$

is satisfied.

Finally, let $t$ lie in the interior of the convex hull of the support of $\psi$.

Let

$$t_{\text{crit}} := \int_{-\infty}^{\infty} y \psi(y) dy$$

denote the expectation of the density $\psi(y)$.

Given $\alpha \in I$, we may define a tilted probability density $\psi_\alpha$ by setting

$$\psi_\alpha(y) := \frac{e^{\alpha y} \psi(y)}{Z(\alpha)}. \quad (29)$$

The following lemma is similar to Lemma 1.3, and says that under Assumption 4.1, there exists a tilted version of the density $\psi$ whose mean is given by $t$. 
Lemma 4.2. Under Assumption 4.1, there exists \( \alpha_t \in (\alpha_{\inf}, \alpha_{\sup}) \) such that \( \log Z(\alpha_t) \)' \( = t \), and hence
\[
\int_{-\infty}^{\infty} y \psi_{\alpha_t}(y) dy = t.
\]

Additionally, if \( t < t_{\text{crit}} \), then \( \alpha_t < 0 \), whereas if \( t > t_{\text{crit}} \), \( \alpha_t > 0 \).

The proof of Lemma 4.2 is given in the Appendix.

Clearly for any tilted measure \( \psi_{\alpha_t} \) as in Lemma 4.2, the moment generating function of \( \psi_{\alpha_t} \) exists in a neighbourhood around zero, and hence \( \psi_{\alpha_t} \) has moments of all orders. Define
\[
m_k(\alpha) := \int_{-\infty}^{\infty} (y - \alpha)^k \psi_{\alpha_t}(y) dy
\]
to be the centered \( k \)-th moment associated with \( \psi_{\alpha_t} \). For \( t \) as in Assumption 4.1 and the associated parameter \( \alpha_t \) as in Lemma 4.2 write
\[
\sigma_t := m_2(\alpha_t), \quad \kappa_t := \frac{m_3(\alpha_t)}{m_2(\alpha_t)^{3/2}} \quad \text{and} \quad R_t := \frac{m_4(\alpha_t)}{m_2(\alpha_t)^2}.
\]

Loosely speaking, Cramér’s theorem \([6, \text{Section 2.2}]\) asserts that whenever \( t < t_{\text{crit}} \) we have
\[
\lim_{N \to \infty} \frac{1}{N} \log P \left( Y_1 + \ldots + Y_N \leq tN \right) = I(t),
\]
where \( I : [t_0, \infty) \to [0, \infty) \) is a rate function given by
\[
I(t) := \alpha_t t - \log Z(\alpha_t),
\]
where \( \alpha_t \) is as in Lemma 4.2.

In the present paper we will appeal to require quantitative versions of Cramér’s theorem. The following result, for which we provide a proof in the Appendix, is a reformulation of well-known results in the literature (see e.g. Bahadur and Ranga Rao \([2]\)).

Lemma 4.3. Suppose under Assumption 4.1 we additionally have \( t < t_{\text{crit}} \). Then we have
\[
P \left( Y_1 + \ldots + Y_N \leq tN \right) = 1 + O \left( \frac{R_t}{\sqrt{N}} \right) \frac{1}{\sqrt{2\pi \sigma_t^2 \alpha_t^2 N}} e^{-NI(t)}.
\]

where the \( O \) term is universal.

We now turn to discussing the Gibbs conditioning principle \([6, \text{Section 7.3}]\). Roughly speaking, this principle asserts that for \( t < t_{\text{crit}} \), conditioned on the event \( \{ Y_1 + \ldots + Y_N \leq tN \} \), the \( k \)-dimensional random vector \( (Y_1, \ldots, Y_k) \) converges in distribution to the \( k \)-dimensional product density \( \psi_{\alpha_t}^{\otimes k} \) as \( N \to \infty \). Shortly we will give a quantitative statement of this principle in terms of total variation distances. Define the constant
\[
\xi := \frac{1}{2} \int_{-\infty}^{\infty} e^{-\frac{1}{2} \xi^2} \left| 1 - \zeta^2 \right| d\zeta = \sqrt{\frac{2}{\pi e}},
\]
where the final equality above follows from noting \( \frac{d}{d\zeta} \left( \zeta e^{-\frac{\zeta^2}{2}} \right) = (1 - \zeta^2) e^{-\frac{\zeta^2}{2}} \). Now for \( \theta \in (0,1) \) define
\[
Q(\theta) := \int_{-\infty}^{\infty} \left| 1 - \sqrt{1 - \theta} e^{\theta \xi^2/2} \right| \frac{1}{\sqrt{2\pi}} \exp \left( -\frac{1}{2} \frac{\xi^2}{\theta^2} \right) d\zeta.
\]

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 \([8]\).
Moreover, we note that by definition
$$\mu = \left(1 + \varepsilon_{k,N}\right) \xi_{k,N}$$
where for given by the conditional law of the first
$$k$$
on $$\phi$$.

**Lemma 4.4** (Theorem 1.6 of [8]). Let $$\psi_{t}^{N \rightarrow k}$$ be the marginal density of the vector $$(Y_1, \ldots, Y_k)$$ conditioned on the event $$\{Y_1 + \ldots + Y_N \leq tN\}$$. Then
$$\int_{\mathbb{R}^k} \left| \psi_{t}^{N \rightarrow k}(y) - \psi_{t}^{\phi}(y) \right| dy = \begin{cases} (1 + \varepsilon_{\phi,N}) \xi_{k,N} & : k = o(N), N \to \infty, \\ (1 + \varepsilon_{\phi,N}) Q(\theta) & : k \sim \theta N, k, N \to \infty, \end{cases}$$
where $$\varepsilon_{k,N}$$ and $$\varepsilon_{\phi,N}$$ are $$O \left( \frac{1}{\sqrt{k}} + \frac{1}{\sqrt{N/k}} \right)$$.

Since our statement is slightly different from [8, Theorem 1.6], a proof of Lemma 4.4 is given in the Appendix.

We now have all the tools at hand to prove Theorem 1.4, which we do over the next two sections.

### 5. Proof of Theorem 1.4: the $$t > t_{\text{crit}}$$ case

Recall that according to the first point in Theorem 1.4, 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
$$\int_{\mathbb{R}^k} \left| \mu_{\phi,t}^{N \rightarrow k}(s) - \gamma_{\phi,\text{uni}}(s) \right| ds \leq C e^{-cN}. \quad (32)$$

As mentioned in the introduction, we in fact prove the following stronger statement, of which the bound (32) is a consequence.

**Theorem 5.1.** If $$t_{\text{crit}} < t < t_{\text{sup}}$$,
$$\int_{\mathbb{R}^k} \left| \mu_{\phi,t}^{N \rightarrow k}(s) - \gamma_{\phi,\text{uni}}(s) \right| ds = (2 + \varepsilon_{N,k}) \frac{1}{\sqrt{2\pi \sigma_t^2 \alpha_t^2}} e^{-I(t)N},$$

where the rate function $$I : [0, t_{\text{sup}}) \to [0, \infty)$$ is given by $$I(t) := t\alpha_t - \log Z(\alpha_t), \sigma_t^2 := \frac{\partial^2}{\partial \alpha^2} \log Z(\alpha) \big|_{\alpha = \alpha_t}$$ and there are constants $$C = C(\phi, t), c = c(\phi, t) \in (0, \infty)$$ such that $$|\varepsilon_{N,k}| \leq C \left( \frac{1}{\sqrt{N}} + \frac{k}{N} + e^{-cN} \right)$$ 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,
$$Q_{\phi,t}^{N \rightarrow k}(s) := \frac{1}{|D_{\phi}^N| - |B_{\phi,t}^N|} \int_{\mathbb{R}^{N-k}} \mathbb{1}_{\{ (s, \zeta) \in D_{\phi}^N \setminus B_{\phi,t}^N \}} d\zeta, \quad (33)$$

where for $$s = (s_1, \ldots, s_k) \in \mathbb{R}^k$$ and $$\zeta = (\zeta_1, \ldots, \zeta_{N-k}) \in \mathbb{R}^{N-k}$$, we write $$(s, \zeta) := (s_1, \ldots, s_k, \zeta_1, \ldots, \zeta_{N-k}) \in \mathbb{R}^N$$. Moreover, we note that by definition $$\mu_{\phi,t}^{N \rightarrow k}$$ may also be written as an integral over $$\mathbb{R}^{N-k}$$:
$$\mu_{\phi,t}^{N \rightarrow k}(s) := \frac{1}{|D_{\phi,t}^N|} \int_{\mathbb{R}^{N-k}} \mathbb{1}_{\{ (s, \zeta) \in D_{\phi,t}^N \}} d\zeta. \quad (34)$$

We now work to express the total variation distance between $$\mu_{\phi,t}^{N \rightarrow k}(s)$$ and $$\gamma_{\phi,\text{uni}}$$ in terms of $$Q_{\phi,t}^{N \rightarrow k}$$. Indeed, by definition we have
$$\int_{\mathbb{R}^k} \left| \mu_{\phi,t}^{N \rightarrow k}(s) - \gamma_{\phi,\text{uni}}(s) \right| ds := \int_{\mathbb{R}^k} \left| \int_{\mathbb{R}^{N-k}} \left( \frac{\mathbb{1}_{\{ (s, \zeta) \in D_{\phi,t}^N \}}}{|D_{\phi,t}^N|} - \frac{\mathbb{1}_{\{ (s, \zeta) \in D_{\phi}^N \}}}{|D_{\phi}^N|} \right) \right| d\zeta \, ds. \quad (35)$$

Now, for each $$s \in \mathbb{R}^k$$, since $$B_{\phi,t}^N \subseteq D_{\phi}^N$$, we have
$$\int_{\mathbb{R}^{N-k}} \left( \frac{\mathbb{1}_{\{ (s, \zeta) \in D_{\phi,t}^N \}}}{|D_{\phi,t}^N|} - \frac{\mathbb{1}_{\{ (s, \zeta) \in D_{\phi}^N \}}}{|D_{\phi}^N|} \right) d\zeta = \left( \frac{1}{|B_{\phi,t}^N|} - \frac{1}{|D_{\phi,t}^N|} \right) \int_{\mathbb{R}^{N-k}} \mathbb{1}_{\{ (s, \zeta) \in D_{\phi,t}^N \}} d\zeta - \frac{1}{|D_{\phi}^N|} \int_{\mathbb{R}^{N-k}} \mathbb{1}_{\{ (s, \zeta) \in D_{\phi,t}^N \}} d\zeta.$$
Similarly, we may also write \( \mu_{\phi,t}^{N-k} \) as
\[
\left(1 - \frac{|B_{\phi,t}^N|}{|D_{\phi}|^N}\right) (\mu_{\phi,t}^{N-k}(s) - Q_{\phi,t}^{N-k}(s))
\]
where we used (33) and (34) to obtain the final equality above. Plugging (36) into (35), we obtain
\[
\int_{\mathbb{R}^k} \left| \mu_{\phi,t}^{N-k}(s) - \gamma_{\phi,uni}^{\otimes k}(s) \right| \, ds = \left(1 - \frac{|B_{\phi,t}^N|}{|D_{\phi}|^N}\right) \int_{\mathbb{R}^k} \left| Q_{\phi,t}^{N-k}(s) - \gamma_{\phi,uni}^{\otimes k}(s) \right| \, ds.
\]

The following lemma is the main part of the proof, giving a fine estimate of the integral occurring in (37).

**Lemma 5.2.** Fix \( \theta \in (0,1) \). There are constants \( C = C(\theta, \phi, t), c = c(\theta, \phi, t) \in (0, \infty) \) such that for all integers \( k, N \) such that \( k \leq \theta N \) we have
\[
\int_{\mathbb{R}^k} \left| Q_{\phi,t}^{N-k}(s) - \gamma_{\phi,uni}^{\otimes k}(s) \right| \, ds = 2 - \varepsilon_{N,k},
\]
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 \( \gamma_{\phi,uni}^{\otimes k} \) 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 \( \Phi : D_{\phi}^k \to [0, \infty) \) is given by \( \Phi(s_1, \ldots, s_k) := (\phi(s_1), \ldots, \phi(s_k)) \) and \( f : [0, \infty)^k \to [0, \infty) \) is given by
\[
f(y_1, \ldots, y_k) := \frac{1}{|D_{\phi}|^k - |B_{\phi,t}^N|} \int_{\mathbb{R}^{N-k}} 1_{\{\phi(1) + \cdots + \phi(N-k) > tN - \sum_{i=1}^k y_i\}} \, d\zeta_1 \cdots d\zeta_{N-k}.
\]
Similarly, we may also write \( \gamma_{\phi,uni}^{\otimes k} := g(\Phi(s)) \), where \( g(y_1, \ldots, y_k) := \frac{1}{|D_{\phi,t}^N|} \) (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
\[
\int_{\mathbb{R}^k} \left| Q_{\phi,t}^{N-k}(s) - \gamma_{\phi,uni}^{\otimes k}(s) \right| \, ds = \int_{[0, \infty)^k} \left| \Phi^* Q_{\phi,t}^{N-k}(y) - \Phi^* \gamma_{\phi,uni}^{\otimes k}(y) \right| \, dy,
\]
where \( \Phi^* Q_{\phi,t}^{N-k} \) and \( \Phi^* \gamma_{\phi,uni}^{\otimes k} \) denote the respective densities on \([0, \infty)^k\) of the random vectors \( (\phi(X_1), \ldots, \phi(X_k)) \) and \( (\phi(Y_1), \ldots, \phi(Y_k)) \) where \( (X_1, \ldots, X_k) \) is distributed according to density \( Q_{\phi,t}^{N-k} \) and \( (Y_1, \ldots, Y_k) \) is distributed according to density \( \gamma_{\phi,uni}^{\otimes k} \).

Now note that \( \Phi^* Q_{\phi,t}^{N-k} \) is precisely the conditional density of \( (\phi(X_1), \ldots, \phi(X_k)) \) conditioned on the event \( \{\phi(1) + \cdots + \phi(N-k) > tN\} \) where \( X_1, \ldots, X_N \) are independent and uniformly distributed on \( D_{\phi} \). Equivalently, by (27), \( \Phi^* Q_{\phi,t}^{N-k} \) is the conditional density of \( (Y_1, \ldots, Y_k) \) conditioned on the event \( \{Y_1 + \cdots + Y_N > tN\} \), where \( Y_1, \ldots, Y_N \) are independent and identically distributed with density \( \psi_{\phi}(y) := \psi(y) / |D_{\phi}| \). In particular, applying Lemma 4.4 to the variables \( -Y_1, \ldots, -Y_k \), and extracting a rather rough bound from Lemma 4.4, we see that there is a constant \( C = C(\phi, t) \in (0, \infty) \) such that for all \( k, N \in \mathbb{N} \) with \( k \leq N \)

\[
\int_{\mathbb{R}^k} \left| \Phi^* Q_{\phi,t}^{N-k}(y) - \psi^{\otimes k}_{\alpha}(y) \right| \, dy \leq C \frac{k}{N}
\]
where \( \alpha \) is chosen so that \( t = \frac{\alpha}{\log Z(\alpha)} \).\( \alpha = \alpha_t. \) We remark that the precision granted by Lemma 4.4 is utilised fully in the next section in our study of the case \( t \leq t_{\text{crit}}. \)

In particular, using (38), (39) and the triangle inequality we have
\[
\int_{\mathbb{R}^k} \left| Q_{\phi,t}^{N-k}(s) - \gamma_{\phi,uni}^{\otimes k}(s) \right| \, ds = \int_{[0, \infty)^k} \left| \Phi^* \gamma_{\phi,uni}^{\otimes k}(y) - \psi^{\otimes k}_{\alpha}(y) \right| \, dy + \Delta_{k,N},
\]
where \( |\Delta_{k,N}| \leq C k / N. \)

We now note that \( \Phi^* \gamma_{\phi,uni}^{\otimes k} = \psi^{\otimes k}_{\alpha} \). In particular,
\[
\int_{[0, \infty)^k} \left| \Phi^* \gamma_{\phi,uni}^{\otimes k}(y) - \psi^{\otimes k}_{\alpha}(y) \right| \, dy = 2 \sup_{A \subseteq [0, \infty)^k} \left( \int_A \psi^{\otimes k}_{\alpha}(y) \, dy - \int_A \psi^{\otimes k}(y) \, dy \right)
\]
is the total variation between the distributions of \((Y_1, \ldots, Y_k)\) and \((Y_1', \ldots, Y'_k)\), 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 \(\psi_{\alpha_i}\), 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 \(\left\{ \frac{Y_1 + \ldots + Y_k}{k} \leq \nu \right\}\) but unlikely that \(\left\{ \frac{Y_1' + \ldots + Y'_k}{k} \leq \nu \right\}\). More explicitly, setting \(A_u := \left\{ (y_1, \ldots, y_k) \in [0, \infty)^k : y_1 + y_2 + \ldots + y_k > uk \right\}\) and extracting a rather rough bound from the quantitative Cramér theorem, Lemma 4.3, we see that there exist constants \(C\) and \(C'\) such that \(\int_{A_u} \psi_{\alpha}^{\otimes k}(y)dy \leq Ce^{-ck}\) and \(\int_{A_u} \psi_{t}^{\otimes k}(y)dy \geq 2 - Ce^{-ck}\).

In particular, two previous two estimates imply that
\[
2 \geq \int_{[0, \infty)^k} \left| \Phi_{\phi, \text{uni}}^{\otimes k}(y) - \Phi_{t, \text{uni}}^{\otimes k}(y) \right|dy \geq 2 - 2Ce^{-ck}.
\]
Combining (40) with (41), 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 (37).

On the one hand, we may write
\[
\left( 1 - \frac{|B_{\phi, t}^N|}{|D_{\phi}|^N} \right) = \mathbb{P} \left[ \phi(V_1) + \ldots + \phi(V_N) > tN \right],
\]
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 Cramér’s theorem, Lemma 4.3, we have
\[
\left( 1 - \frac{|B_{\phi, t}^N|}{|D_{\phi}|^N} \right) = \left( 1 + \varepsilon_N \right) \frac{1}{\sqrt{2\pi \sigma_\phi^2 \alpha_\phi^2 N}} \exp\left( -ItN \right),
\]
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
\[
\int_{\mathbb{R}^k} \left| Q_{\phi, \text{uni}}^{N+k}(s) - \gamma_{\phi, \text{uni}}^{\otimes k}(s) \right|ds = 2 - \varepsilon_{N,k},
\]
where for a different constant \(C'' = C''(\phi)\) we have \(|\varepsilon_{N,k}| \leq C'' \left( \frac{k}{N} + e^{-ck} \right)\).

In particular, by (37),
\[
\int_{\mathbb{R}^k} \left| \Phi_{\phi, t}^{N+k}(s) - \gamma_{\phi, \text{uni}}^{\otimes k}(s) \right|ds = \left( 1 + \varepsilon_N \right) \left( 2 - \varepsilon_{N,k} \right) \frac{1}{\sqrt{2\pi \sigma_\phi^2 \alpha_\phi^2 N}} \exp\left( -ItN \right),
\]
where \(|\varepsilon_{N,k}| \leq C'' \left( \frac{k}{N} + e^{-ck} \right)\) and \(|\varepsilon_N| < C/\sqrt{N}\). Set \(\rho_{N,k}\) to be the solution to
\[
1 + \rho_{N,k} = \left( 1 + \varepsilon_N \right) \left( 1 - \frac{1}{2} \varepsilon_{N,k} \right)
\]
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 1.4: the $t \leq t_{crit}$ case

6.1. A full statement and overview

We now turn to proving Theorem 1.4 in the case where $t \leq t_{crit}$. 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

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

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

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

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

$$\int_{\mathbb{R}^k} |\mu_{\phi,t}^{N \to k}(x) - \gamma_{\phi, \alpha}(x)| \, dx = \begin{cases} (1 + \varepsilon_{k,N}) \frac{\xi_k}{N} & : k = o(N), N \to \infty, \\ (1 + \varepsilon'_{k,N}) Q(\theta) & : k \sim \theta N, k, N \to \infty, \end{cases}$$

where $\varepsilon_{k,N}$ and $\varepsilon'_{k,N}$ are $O\left( R_t \left( \frac{1}{\sqrt{k}} + \frac{1}{\sqrt{N-k}} \right) \right)$.

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 $\Phi_{\alpha_t}^N \mu_{\phi,t}^{N \to k}$ and $\Phi_{\alpha_t}^N \gamma_{\phi, \alpha}^{\otimes k}$.
- **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 $\mu_{\phi,t}^{N \to k}$ and $\gamma_{\phi, \alpha}^{ \otimes k}$.
- **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 $t \leq t_{crit}$. Then, with $\alpha_t \in \mathbb{R}$ as in Definition 1.3, we have

$$\int_{\mathbb{R}^k} |\Phi_{\alpha_t}^N \mu_{\phi,t}^{N \to k}(y) - \psi_{\alpha_t} \gamma_{\phi, \alpha}(y)| \, dy = \begin{cases} (1 + \varepsilon_{k,N}) \frac{\xi_k}{N} & : k = o(N), N \to \infty, \\ (1 + \varepsilon'_{k,N}) Q(\theta) & : k \sim \theta N, k, N \to \infty, \end{cases}$$

where $\varepsilon_{k,N}$ and $\varepsilon'_{k,N}$ are $O\left( R_t \left( \frac{1}{\sqrt{k}} + \frac{1}{\sqrt{N-k}} \right) \right)$.

**Proof.** Since $|D_{\phi}| < \infty$, the uniform density $\gamma_{\phi, uni}(s) := \frac{1_{D_{\phi}}(s)}{|D_{\phi}|}$ on $D_{\phi}$ exists. Suppose now $X_1, \ldots, X_N$ are independent random variables distributed according to the uniform density $\gamma_{\phi, uni}$ on $D_{\phi}$, and consider the transformed variables $\phi(Y_1), \ldots, \phi(Y_N)$, which are distributed according to $\psi_0$, where $\psi_0$ is given in (27). Noting that by definition $\mu_{\phi,t}^{N \to k}$ is the marginal density of $(X_1, \ldots, X_k)$ conditioned on the event $\{\phi(Y_1) + \cdots + \phi(Y_N) \leq tN\}$, it follows that the pushforward $\Phi_{\alpha_t}^N \mu_{\phi,t}^{N \to k}$ is the marginal density of $(Y_1, \ldots, Y_k)$ conditioned on the event $\{Y_1 + \cdots + Y_N \leq tN\}$. In other words, we are in the setting of Lemma 4.4, from which the result follows immediately.

6.3. Step 2

The next lemma ‘pulls back’ the previous result, replacing the estimate of the total variation between $\Phi_{\alpha_t}^N \mu_{\phi,t}^{N \to k}$ and $\psi_{\alpha_t} \gamma_{\phi, \alpha}^{ \otimes k}$ with one between $\mu_{\phi,t}^{N \to k}$ and $\gamma_{\phi, \alpha}^{ \otimes k}$. 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 $\alpha_t \in \mathbb{R}$ as in Definition 1.3, we have

$$
\int_{\mathbb{R}^k} \left| \mu_{\phi,t}^N - \gamma_{\phi,t}^k(s) \right| ds = \begin{cases} (1 + \varepsilon_{k,N}) \xi_N k & : k = o(N), N \to \infty, \\ (1 + \varepsilon_{k,N}') Q(\theta) & : k \sim \theta N, k, N \to \infty, 
\end{cases}
$$

where $\varepsilon_{k,N}$ and $\varepsilon_{k,N}'$ are $O \left( \frac{1}{\sqrt{k}} + \frac{1}{\sqrt{N-k}} \right)$.

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

$$
f(y_1, \ldots, y_k) := \frac{1}{|D_{\phi,t}^N|} \int_{\mathbb{R}^{N-k}} \mathbb{1}_{\{\phi(s_1) + \cdots + \phi(s_{N-k}) \leq \sum_{i=1}^k y_i\}} d\zeta_1 \cdots d\zeta_{N-k}.
$$

Moreover, clearly $\gamma_{\phi,t}^k(s) = g(\Phi(s))$, where $g(y_1, \ldots, y_k) = \frac{\sum_{i=1}^k y_i}{Z(\alpha_t)}$. In particular, we are in the setting of Lemma 3.5, so that

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

The result follows by noting that $\Phi^N_{\gamma_{\phi,t}^k} = \Phi^N\gamma_{\phi,t}^k$, and using Lemma 6.2. \qed

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

$$
\phi_L(s) := \begin{cases} \phi(s) & \text{if } s \leq L \\
\infty & \text{if } s > L.
\end{cases}
$$

Define the truncated partition function

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

where the final equality above is a truncated analogue of (25), and follows from an application of (24). For $t \in \mathbb{R}$ let $\alpha_{t,L}$ denote the solution to $\frac{\partial}{\partial t} Z_L(\alpha)|_{\alpha = \alpha_{t,L}} = t$.

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

$$
||\mu_{\phi,t}^N - \gamma_{\phi,t}^k|| \leq ||\mu_{\phi_{L,t}}^N - \gamma_{\phi_{L,t}}^k|| + ||\mu_{\phi_{L,t}}^N - \mu_{\phi,t}^N|| + ||\gamma_{\phi_{L,t},\alpha_{t,L}} - \gamma_{\phi,t}^k||
$$

and

$$
||\mu_{\phi,t}^N - \gamma_{\phi,t}^k|| \leq ||\mu_{\phi_{L,t}}^N - \gamma_{\phi_{L,t}}^k|| + ||\mu_{\phi_{L,t}}^N - \gamma_{\phi_{L,t}}^k|| + ||\gamma_{\phi_{L,t},\alpha_{t,L}} - \gamma_{\phi,t}^k||
$$

In particular, we may write

$$
\int_{\mathbb{R}^k} \left| \mu_{\phi,t}^N - \gamma_{\phi,t}^k(s) \right| ds = \int_{\mathbb{R}^k} \left| \mu_{\phi_{L,t}}^N - \gamma_{\phi_{L,t}}^k(s) \right| ds + \Delta_L,
$$

(42)
where

\[ |\Delta_L| \leq \int_{\mathbb{R}^k} |\gamma_{\phi,\alpha,L}^{\otimes k}(s) - \gamma_{\phi,\alpha,L}^{\otimes k}(s)| ds + \int_{\mathbb{R}^k} |\mu_{\phi,t,L}^{N\rightarrow k}(s) - \mu_{\phi,t}^{N\rightarrow k}(s)| ds. \quad (43) \]

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 \( B_{\phi,L}^N \) is a subset of \( B_{\phi,L}^N \), both of which have finite Lebesgue measure. Moreover, as \( L \rightarrow \infty \), \( B_{\phi,L}^N \rightarrow \emptyset \), so that by the continuity of Lebesgue measure we have \( |B_{\phi,L}^N \setminus B_{\phi,L}^N| \rightarrow 0 \). In particular, by Lemma 3.4 there is an \( L_1 \) such that whenever \( L \geq L_1 \), we have

\[ \int_{\mathbb{R}^k} |\mu_{\phi,t,L}^{N\rightarrow k}(s) - \mu_{\phi,t}^{N\rightarrow k}(s)| ds \leq \rho/2. \quad (44) \]

On the other hand, using Lemma 3.2 we have

\[ \int_{\mathbb{R}^k} |\gamma_{\phi,\alpha,L}^{\otimes k}(s) - \gamma_{\phi,\alpha,L}^{\otimes k}(s)| ds \leq k \int_{\mathbb{R}} |\gamma_{\phi,\alpha,L}(s) - \gamma_{\phi,\alpha,L}(s)| ds. \quad (45) \]

Now consider that \( \gamma_{\phi,\alpha}(s) = \frac{e^{\alpha(s)}}{Z(\alpha)} \) and \( \gamma_{\phi,\alpha,L}(s) = \frac{e^{\alpha_L(s)}}{Z_L(\alpha)} \) 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

\[ \int_{\mathbb{R}} |\gamma_{\phi,\alpha,L}(s) - \gamma_{\phi,\alpha,L}(s)| ds = \int_{0}^{\infty} |\phi' \gamma_{\phi,\alpha,L}(y) - \phi' \gamma_{\phi,\alpha,L}(y)| dy. \quad (46) \]

We now note that in the notation of Section 2, we have \( \phi' \gamma_{\phi,\alpha,L}(y) = \left( \phi' \gamma_{\phi,\alpha,L}(y) \right)_L \), so that making the relevant substitution in (46) we have

\[ \int_{\mathbb{R}} |\gamma_{\phi,\alpha,L}(s) - \gamma_{\phi,\alpha,L}(s)| ds = \int_{0}^{\infty} \left| \left( \phi' \gamma_{\phi,\alpha,L}(y) - \phi' \gamma_{\phi,\alpha,L}(y) \right)_L \right| dy. \]

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

\[ \int_{0}^{\infty} \left| \left( \phi' \gamma_{\phi,\alpha,L}(y) - \phi' \gamma_{\phi,\alpha,L}(y) \right)_L \right| dy \leq \rho/2k. \quad (47) \]

In particular, combining (47) and (45) with (44) in (43), we see that for all \( L \geq L_0 := \max\{L_1, L_2\} \),

\[ |\Delta_L| \leq \rho. \quad (48) \]

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

\[ \int_{\mathbb{R}^k} |\mu_{\phi,t,L}^{N\rightarrow k}(s) - \mu_{\phi,t,L}^{\otimes k}(s)| ds \begin{cases} (1 + \varepsilon_{k,N}) \xi_{k,N} & : k = o(N), N \rightarrow \infty, \\ \left(1 + \varepsilon_{k,N}\right) Q(\theta) & : k \sim \theta N, k, N \rightarrow \infty, \end{cases} \quad (49) \]

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

\[ \left| \varepsilon_{k,N}, \varepsilon_{k,N}' \right| \leq C'_L \left( \frac{1}{\sqrt{k}} + \frac{1}{\sqrt{N-k}} \right). \]

Now by applying the third point of Lemma 2.2 to the truncated density \( (\gamma_{\alpha,L})_L \), 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

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

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

\[ \left| \varepsilon_{k,N,L}^{(1)} \right| \leq C' \sqrt{\frac{k}{N}}, \quad \left| \varepsilon_{k,N,L}^{(2)} \right|, \left| \varepsilon_{k,N,L}^{(3)} \right| \leq C' \left( \frac{1}{\sqrt{k}} + \frac{1}{\sqrt{N-k}} \right). \quad (50) \]

In particular, since \( \rho > 0 \) and \( L \) are arbitrary, by combining (42), (48), (49) and (50) we obtain the result. \( \square \)
Appendix

In this appendix we prove various results stated in the main body of the article. In all cases, these results and their proofs are variations on well-known work in the literature.

Proof of Lemma 1.3. We define the auxiliary function $Z(\alpha, t) := Z(\alpha)e^{-\alpha t} = \int_{D_\alpha} e^{\alpha(\phi(s)-t)}ds$, and write $Z'(\alpha, t) = \int_{D_\alpha} (\phi(s) - t)e^{\alpha(\phi(s)-t)}ds$ for its derivative with respect to $\alpha$.

It is easily verified that $Z'(\alpha, t) = 0$ is equivalent to $(\log Z)'(\alpha) = t$. Now since $Z''(\alpha, t) = \int_{D_\alpha} (\phi(s) - t)^2 e^{\alpha(\phi(s)-t)}ds > 0$, and the level sets of $\phi$ have measure zero, it follows that $Z'(\alpha, t)$ is a strictly increasing function of $\alpha$.

Using the fact that the essential infimum and supremum of $\phi$ on its domain $D_\alpha$ are given by 0 and $t_{sup}$ respectively, we now verify that for each $t \in (0, t_{sup})$ we have

$$\lim_{\alpha \downarrow -\infty} Z'(\alpha, t) = -\infty \quad \text{and} \quad \lim_{\alpha \uparrow t_{sup}} Z'(\alpha, t) = \infty. \quad (51)$$

To see that the $\alpha \downarrow -\infty$ limit in (51) holds, consider that we may write

$$Z'(\alpha, t) := \int_{-\infty}^{\infty} (\phi(s) - t)e^{\alpha(\phi(s)-t)}ds,$$

so that $\lim_{\alpha \downarrow -\infty} Z'(\alpha, t) = -\infty$ by the monotone convergence theorem. As for the $\alpha \uparrow t_{sup}$ limit in (51), we note that $\lim_{\alpha \uparrow t_{sup}} Z(\alpha) = \infty$ implies $\lim_{\alpha \uparrow t_{sup}} Z'(\alpha, t) = \infty$, which in turn implies $\lim_{\alpha \uparrow t_{sup}} Z'(\alpha, t) = \infty$.

Since $Z'(\alpha, t)$ is a strictly increasing bijection from $-\infty$ to $+\infty$ as $\alpha$ increases from $-\infty$ to $+\infty$, it follows that for each $t \in (0, t_{sup})$ there is a unique $\alpha_t$ such that $Z'(\alpha_t, t) = 0$, so that $(\log Z)'(\alpha_t) = t$.

It remains to show that $\alpha_t$ has the same sign as $t - t_{crit}$. To see this, we note that $Z'(0, t) = \int_{D_\alpha} (\phi(s) - t)ds = t_{crit} - t$, and hence since $Z'(\alpha, t)$ is increasing in $\alpha$, the sign of the solution $\alpha_t$ to $Z'(\alpha, t) = 0$ is the same as that of $(t - t_{crit})$. \Box

Proof of Lemma 4.2. The proof of Lemma 4.2 runs identically to the proof of Lemma 1.3, though this time instead of the auxiliary function $Z(\alpha, t) := Z(\alpha)e^{-\alpha t} = \int_{-\infty}^{\infty} e^{\alpha(y-t)}\psi(y)dy$.

In the proofs of Lemmas 4.3 and 4.4, we will use the Edgeworth expansion. Suppose that $X_1, X_2, \ldots$ are independent and identically distributed random variables with a finite fourth moment. Let $\mu := \mathbb{E}[X_1], \sigma := \mathbb{E}[(X - \mu)^2]$, and let $g_j$ be the density function of the centered unit variance random variable

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

Then the Edgeworth expansion states that

$$g_j(s) = \frac{1}{\sqrt{2\pi}}e^{-s^2/2} \left(1 + \frac{\mathbb{E}[(X - \mu)^4]}{\mathbb{E}[(X - \mu)^2]^3} \frac{1}{\sqrt{2}}(s^3 - 3s) + O\left(\frac{\mathbb{E}[(X - \mu)^4]}{\mathbb{E}[(X - \mu)^2]^2} s^3\right)\right), \quad (52)$$

where the $O$ term is universal, and uniform in $s$. See e.g. Feller [9, Chapter XVI].

Proof of Lemma 4.3. Let $Y_1, \ldots, Y_N$ be independent and identically distributed with density function $\psi$, and let $Y'_1, \ldots, Y'_N$ be independent and identically distributed with density function $\psi_{\alpha_t}(s) = e^{\alpha_t \psi(s)}Z(\alpha_t)$, where $\alpha_t$ as in Lemma 1.3 (so that the expectation of $\psi_{\alpha_t}$ is $t$). We can then write

$$\mathbb{P}(Y_1 + \ldots + Y_N \leq tN) = \int_{\mathbb{R}^N} 1\{y_1 + \ldots + y_N \leq tN\} \prod_{i=1}^{N} \psi(y_i)dy_i$$

$$= Z(\alpha_t)^N \int_{\mathbb{R}^N} e^{-\alpha_t(y_1 + \ldots + y_N)} 1\{y_1 + \ldots + y_N \leq tN\} \prod_{i=1}^{N} \psi_{\alpha_t}(y_i)dy_i$$

$$= Z(\alpha_t)^N \mathbb{E}[e^{-\alpha_t(Y'_1 + \ldots + Y'_N)} 1\{Y'_1 + \ldots + Y'_N \leq tN\}], \quad (53)$$

$$= Z(\alpha_t)^N \mathbb{E}[e^{-\alpha_t(Y'_1 + \ldots + Y'_N)} 1\{Y'_1 + \ldots + Y'_N \leq tN\}], \quad (54)$$
Defining the zero-mean, unit-variance random variable $S_N := (Y_1' + \ldots + Y_N' - tN)/\sigma_1\sqrt{N}$, this reduces to
\[
\mathbb{P}(Y_1 + \ldots + Y_N \leq tN) := e^{-NI(t)}\mathbb{E}[e^{-\alpha(t)\sqrt{N}\mathbb{Q}} 1_{\mathbb{S}_N \leq 0}],
\] (55)
where the rate function $I(t)$ is defined in (30). Now according to the Edgeworth expansion (52), the density function $g_N(s)$ of $S_N$ may be written
\[
g_N(s) = \frac{1}{\sqrt{2\pi}}e^{-s^2/2} + O\left(\frac{R_t}{\sqrt{N}}\right),
\] (56)
where $\kappa_t$ and $R_t$ are defined in (29), and again, the $O$ term is uniform in $s$ (and universal). Now using Jensen’s inequality and the definitions of $\kappa_t$ and $R_t$, it is straightforward to see both that $R_t \geq 1$ and $\kappa_t \leq R_t^{3/4}$. Consequently $\kappa_t \leq R_t$, and we take a rather rough bound on the expansion (56) to obtain
\[
g_N(s) = \frac{1}{\sqrt{2\pi}}e^{-s^2/2} + O\left(\frac{R_t}{\sqrt{N}}\right),
\] (57)
uniformly in $s$. Plugging (57) into (55) to obtain the first equality below, and then recalling that $\kappa_t < 0$ and integrating to obtain the second, we see that
\[
\mathbb{P}(Y_1 + \ldots + Y_N \leq tN) := e^{-NI(t)}\int_{-\infty}^{0} e^{-\alpha(t)\sqrt{N}s} \left(\frac{1}{\sqrt{2\pi}}e^{-s^2/2} + O\left(\frac{R_t}{\sqrt{N}}\right)\right) ds
\]
\[
= e^{-NI(t)}\left(\frac{1}{\sqrt{2\pi}|\alpha_t|\sqrt{N}} + O(R_t/N)\right),
\]
completing the proof.

Proof of Lemma 4.4. The proof runs almost identically to Section 3 of Diaconis and Freedman [8], which in turn appeals to definitions from Section 2 of their previous work [7]. We will import freely results from the aforementioned sections of [8] and [7].

Consider the probability measures $\mathbb{P}$ and $\mathbb{Q}$ on $\mathbb{R}^k$ with respective densities $\psi_{\leq k}^N$ and $\psi_{\leq k}^*$. Let $\mathcal{F}$ denote the Borel subsets of $\mathbb{R}^k$, and consider the sub-sigma-algebra $\Sigma$ of $\mathcal{F}$ generated by the random variable $X_1 + \ldots + X_k$ summing the coordinates. Then $\Sigma$ is sufficient for $\mathbb{P}$ and $\mathbb{Q}$, in that for $A$ in $\mathcal{F}$ we have the equality
\[
\mathbb{P}(A|\Sigma) = \mathbb{Q}(A|\Sigma).
\]
To see this sufficiency, we note that conditioning on the event $\{X_1 + \ldots + X_k = y\}$, that both under $\mathbb{P}$ and $\mathbb{Q}$ the conditional density of the first $k - 1$ coordinates of the vector $(X_1, \ldots, X_k)$ is proportional to $\psi(x_1)\ldots\psi(x_{k-1})\psi(y-x_1-\ldots-x_{k-1})$, and the final coordinate is determined by the first $k - 1$.

Now, whenever $\mu$ is a probability density on $\mathbb{R}^k$ of a random variable $(X_1, \ldots, X_k)$, let $\bar{\mu}$ denote the probability density on $\mathbb{R}^1$ of $X_1 + \ldots + X_k$. Now according to the sufficiency Lemma 2.4 of [7], we have the equality in total variation
\[
||\psi_{\leq k}^N - \psi_{\leq k}^*|| = ||\psi_{\leq k}^N - \psi_{\leq k}^{\otimes k}||,
\]
where, in the $\bar{\mu}$ notation, $\psi_{\leq k}^N$ refers to the density of $Y_1 + \ldots + Y_k$ conditional on the event $\{Y_1 + \ldots + Y_N \leq tN\}$, and $\psi_{\leq k}^{\otimes k}$ is the density of $Y_1 + \ldots + Y_k$ given independent random variables $Y_1, \ldots, Y_k$ each with density $\psi_{\leq k}^*$.

Write $f^k(s) := \psi_{\leq k}^{\otimes k}(s)$. We remark that if $\psi^k(s)$ is the $k$-fold convolution of the density $\psi$, then
\[
f^k(s) = \frac{e^{\alpha s} \psi^k(s)}{Z(\alpha) s^k}.
\] (58)
Using the definition of $\psi_{\leq k}^N$ to obtain the first equality below, and the definition (58) of $f^k$ to obtain the second, we see that the density associated with $\psi_{\leq k}^N$ may be written
\[
\psi_{\leq k}^N(s) = \frac{\psi^k(s) \int_{-\infty}^{tN-s} \psi^N(u) du}{\int_{-\infty}^{tN} \psi^N(u) du} = \frac{f^k(s) \int_{-\infty}^{tN-s} e^{-\alpha s(u-tN+s)} f_{N-k}(u) du}{\int_{-\infty}^{tN} e^{-\alpha s(u-tN)} f_N(u) du}.
\]
In particular, we have
\[
\| \bar{\psi}^{N-k}_{\leq t} - \bar{\psi}^k_{\alpha_t} \| = \int_{-\infty}^{\infty} \left| f_k(s) \left[ \int_{-\infty}^{t} e^{-\alpha_t(u-tN+s)} f_{N-k}(u) \, du \right] \frac{f_{N}(u)}{f_{N}(u)} - 1 \right| \, ds.
\]

We now consider a centering. Recall \( \sigma_t \) (given in (29)) is the variance of \( \psi_{\alpha_t} \). Given independent random variables \( Y_1, Y_2, \ldots \) identically distributed according to \( \psi_{\alpha_t} \), let \( g_j \) denote the density of
\[
Y_1 + \ldots + Y_j - j t \quad \sigma_t \sqrt{j}
\]
so that when \( j \) is large, \( g_j \) is close to the Gaussian density. Indeed, changing variable we have
\[
\| \bar{\psi}^{N-k}_{\leq t} - \bar{\psi}^k_{\alpha_t} \| = \int_{-\infty}^{\infty} g_k(s) \left| \int_{-\infty}^{t} e^{-\alpha_t \sqrt{N-k}(u+\rho s)} g_{N-k}(u) \, du \frac{f_{N}(u)}{f_{N}(u)} - 1 \right| \, ds,
\]
where \( \rho = \sqrt{\frac{k}{N-k}} \). We now use the Edgeworth expansion to estimate the ratio
\[
\frac{\int_{-\infty}^{t} e^{-\alpha_t \sqrt{N-k}(u+\rho s)} g_{N-k}(u) \, du}{\int_{-\infty}^{t} e^{-\alpha_t \sqrt{N}u} g_{N}(u) \, du}
\]
as a function of the external variable \( s \).

Now considering first the numerator in (60), a calculation using (52) and the moment ratios in (29) tells us that

\[
\int_{-\infty}^{t} e^{-\alpha_t \sqrt{N-k}(u+\rho s)} g_{N-k}(u) \, du = \frac{1}{\sqrt{2\pi} |\alpha_t| \sqrt{N-k}} \left( e^{-(\rho s)^2/2} \left( 1 + \frac{\kappa_t}{\sqrt{N-k}} ((\rho s)^2 - 3\rho s) \right) + O \left( \frac{\rho s^2 + R_t}{\sqrt{N-k}} \right) \right),
\]
where, as above, \( \rho := \sqrt{\frac{k}{N-k}} \).

As for the denominator, again using (52), the moment ratios in (29), and Jensen’s inequality to note that \( \kappa(\alpha_t) = O(R_t) \), we have

\[
\int_{-\infty}^{t} e^{-\alpha_t \sqrt{N}u} g_{N}(u) \, du = \frac{1}{\sqrt{2\pi} |\alpha_t| \sqrt{N}} + O \left( \frac{R_t}{N^{3/2}} \right).
\]

We now use (61) and (62) to appraise the ratio in (60), taking care to distinguish between the cases \( k \sim N \) and \( k = o(N) \). Considering first the case where \( k \sim N \), using (61) and (62) in (60), we have

\[
\frac{\int_{-\infty}^{t} e^{-\alpha_t \sqrt{N-k}(u+\rho s)} g_{N-k}(u) \, du}{\int_{-\infty}^{t} e^{-\alpha_t \sqrt{N}u} g_{N}(u) \, du} = \frac{1}{\sqrt{1-k/N}} \exp \left( -\frac{1}{2} \frac{k}{N-k} s^2 \right) + O \left( \frac{R_t}{\sqrt{N-k}} \right).
\]

By plugging (63) into (59) and setting \( \theta := k/N \) we have

\[
\| \bar{\psi}^{N-k}_{\leq t} - \bar{\psi}^k_{\alpha_t} \| = \int_{-\infty}^{\infty} g_k(s) \left| \frac{1}{\sqrt{1-\theta}} \exp \left( -\frac{1}{2} \frac{\theta}{1-\theta} s^2 \right) - 1 \right| \, ds + O \left( \frac{R_t}{\sqrt{N-k}} \right).
\]

Using the edgeworth expansion (52) to handle the \( g_k(s) \) term in (64), (and noting by Jensen’s inequality we have \( \kappa_t = O(R_t) \), we see that

\[
\| \bar{\psi}^{N-k}_{\leq t} - \bar{\psi}^k_{\alpha_t} \| = \int_{-\infty}^{\infty} e^{-s^2/2} \left| \frac{1}{\sqrt{1-\theta}} \exp \left( -\frac{1}{2} \frac{\theta}{1-\theta} s^2 \right) - 1 \right| \, ds + O \left( \frac{R_t}{\sqrt{N-k}} + \frac{R_t}{\sqrt{N-k}} \right),
\]
completing the proof of Lemma 4.4 in the case where \( k \sim N \).
We now turn to the case where \( k = o(N) \). Here, the ratio takes on the form
\[
\frac{\int_{-\infty}^{-p_0} e^{-\alpha u \sqrt{N - k(u + p_0)}} g_{N-k}(u) \, du}{\int_{0}^{\infty} e^{-\alpha u \sqrt{N}} g_{N}(u) \, du} = 1 + \frac{1}{2N} \frac{k}{N} (1 - s^2) + O \left( \frac{R_t \sqrt{k}}{N} \right).
\] (66)

By plugging (66) into (59), when \( k = o(N) \) we have and setting \( \theta \) we have
\[
\|\psi_{N \leq k} - \psi_{k} \| = k \frac{2N}{N} \int_{-\infty}^{\infty} g_k(s) \left| 1 - s^2 \right| \, ds + O \left( \frac{R_t \sqrt{k}}{N} \right).
\] (67)

Like above, using the edgeworth expansion (52) to handle the \( g_k(s) \) term in (67), we have
\[
\|\psi_{N \leq k} - \psi_{k} \| = k \frac{2N}{N} \int_{-\infty}^{\infty} \frac{e^{-s^2/2}}{\sqrt{2\pi}} \left| 1 - s^2 \right| \, ds + O \left( \frac{R_t \sqrt{k}}{N} \right),
\] (68)
completing the proof in the case \( k = o(N) \). \( \Box \)

Acknowledgements

Both authors have have been supported by the Austrian Science Fund (FWF) Project P32405 “Asymptotic Geometric Analysis and Applications” of which IP is principal investigator.

The second author is also supported by the Special Research Program Project F5508-N26.

References

[1] Alonso-Gutiérrez, D. and Prochno, J. (2022). Thin-shell concentration for random vectors in Orlicz balls via moderate deviations and Gibbs measures. J. Funct. Anal. 282 109291.
[2] Bahadur, R. R. and Rao, R. R. (1960). On deviations of the sample mean. Ann. Math. Statist. 31 1015–1027.
[3] Blum, J. R. and Pathak, K. P. (1972). A note on the zero-one law. Ann. Math. Statist. 43 1008-1009.
[4] Borel, E. (1941). Introduction géométrique à quelques théories physiques. Cornell University Library historical math monographs. Gauthier-Villars.
[5] Daniels, H. E. (1954). Saddlepoint approximations in statistics. Ann. Math. Statist. 25 631–650. MR66602
[6] Dembo, A. and Zeitouni, O. (2010). Large Deviations Techniques and Applications, second edition. Stochastic Modelling and Applied Probability. Springer-Verlag Berlin Heidelberg. MR2571413
[7] Diaconis, P. and Freedman, D. (1987). A dozen de Finetti-style results in search of a theory. Ann. Inst. H. Poincaré Probab. Statist. 23 397–423. MR98502
[8] Diaconis, P. and Freedman, D. A. (1988). Conditional limit theorems for exponential families and finite versions of de Finetti’s theorem. J. Theoret. Probab. 1 381–410. MR958245
[9] Feller, W. (1971). An introduction to probability theory and its applications. Vol. II. Second edition. John Wiley & Sons, Inc., New York-London-Sydney. MR0270403
[10] Fruehwirth, L. and Prochno, J. (2021). Sanov-type large deviations and conditional limit theorems for high-dimensional Orlicz balls. arXiv preprint arXiv:2111.04691.
[11] Horrmann, J., Prochno, J. and Thäle, C. (2018). On the isotropic constant of random polytopes with vertices on an \( \ell_p \)-sphere. J. Geom. Anal. 28 405–426. MR3745865
[12] Kabluchko, Z. and Prochno, J. (2021). The maximum entropy principle and volumetric properties of Orlicz balls. J. Math. Anal. Appl. 495 124687. MR4172842
[13] Kabluchko, Z., Prochno, J. and Thäle, C. (2019). High-dimensional limit theorems for random vectors in \( \ell_p^n \)-balls. Commun. Contemp. Math. 21 1750092, 30. MR3904638
[14] Kabluchko, Z., Prochno, J. and Vysotsky, V. (2020). Yet another note on the arithmetic-geometric mean inequality. Studia Math. 253 39–55. MR4074400
[15] Kim, S. S., Liao, Y. T. and Ramanan, K. (2019). An asymptotic thin shell condition and large deviations for random multidimensional projections. arXiv e-prints arXiv:1912.13447.
[16] Kim, S. S. and Ramanan, K. (2018). A conditional limit theorem for high-dimensional \( \ell_p \)-spheres. J. Appl. Probab. 55 1060–1077. MR3899928
[17] Liao, Y. T. and Ramanan, K. (2020). Geometric sharp large deviations for random projections of \( \ell_p^n \) spheres and balls. arXiv preprint arXiv:2001.04053.
[18] Naor, A. and Rörik, D. (2003). Projecting the surface measure of the sphere of \( \ell_p^n \). Ann. Inst. H. Poincaré Probab. Statist. 39 241–261. MR1962135
[19] Prochno, J., Thäle, C. and Turchi, N. (2019). The isotropic constant of random polytopes with vertices on convex surfaces. J. Complexity 54 101394, 17. MR3983208
[20] Rachev, S. T. and Rüschendorf, L. (1991). Approximate independence of distributions on spheres and their stability properties. Ann. Probab. 19 1311–1337. MR1112418
[21] Schechtman, G. and Zinn, J. (1990). On the volume of the intersection of two \( \ell_p^n \) balls. Proc. Amer. Math. Soc. 110 217–224. MR1015684