

On the Subadditivity of the Entropy on the Sphere

Amit Einav¹

Received: 23 November 2014 / Published online: 17 November 2015
© Mathematica Josephina, Inc. 2015

Abstract We present a refinement of a known entropic inequality on the sphere, finding suitable conditions under which the uniform probability measure on the sphere behaves asymptotically like the Gaussian measure on \mathbb{R}^N with respect to the entropy. Additionally, we remark about the connection between this inequality and the investigation of the many-body Cercignani's conjecture.

Keywords Equivalence of ensembles · Entropy · Entropic inequalities · Transportation distances · Kac's model

Mathematics Subject Classification 39B62 · 52A40 · 60D99 · 82C40 · 94A17

1 Introduction

A fundamental principle in equilibrium statistical mechanics is that of the equivalence of ensembles. In mathematical terms, this principle states that the uniform measure on $\mathbb{S}^{N-1}(\sqrt{N})$, $d\sigma^N$, considered as a measure on \mathbb{R}^N supported on the sphere, is close in behavior to the Gaussian measure

$$d\gamma_N(v) = \frac{e^{-\frac{|v|^2}{2}}}{(2\pi)^{\frac{N}{2}}} dv$$

when N is very large. In this setting the uniform measure $d\sigma^N$ corresponds to the micro-canonical ensemble, representing a fixed number of particles with a fixed total

✉ Amit Einav
a.einav@dpmmms.cam.ac.uk

¹ University of Cambridge, Cambridge, UK

energy, while the Gaussian measure $d\gamma_N$ corresponds to the canonical ensemble, representing a fixed number of particles in thermal equilibrium. For simple systems, the equivalence of ensembles principle means that for any finitely many number of particles with velocities $v_1, \dots, v_k, k \in \mathbb{N}$, and any observable function of those particles, $\phi(v_1, \dots, v_k)$, the measurement of ϕ in the micro-canonical and canonical settings yields almost identical results, with a difference that converges to zero as the number of particles goes to infinity. In other words:

$$\lim_{N \rightarrow \infty} \left(\int_{\mathbb{S}^{N-1}(\sqrt{N})} \phi(v_1, \dots, v_k) d\sigma^N - \int_{\mathbb{R}^N} \phi(v_1, \dots, v_k) d\gamma_N \right) = 0.$$

An acute difference between $d\sigma^N$ and $d\gamma_N$ may arise when one deals with quantities that depend on *all* the particles in the ensemble, such as the case of the entropy, or more generally—the relative entropy, in non-equilibrium statistic mechanics. Such a deviation from the equivalence of ensembles principle was observed in [3]. Before we delve into it, we remind the reader of a few conventions and definitions, so that the work presented here will be self contained.

In what follows, we denote by $P(X)$ the set of Borel probability measures on a Polish space X . Any measure in this current work will be assumed to be a Borel measure.

Definition 1.1 Let $\mu, \nu \in P(\mathbb{R}^d)$. The relative entropy of μ with respect to ν is defined as

$$H(\mu|\nu) = \begin{cases} \int_{\mathbb{R}^d} h \log h d\nu & h = \frac{d\mu}{d\nu} \\ \infty & \text{otherwise.} \end{cases}$$

Note that we have not indicated the dimension of the underlying space in the notation of the relative entropy. It will be implicitly evident in all our discussions to follow.

Definition 1.2 Let $\mu \in P(\mathbb{S}^{N-1}(\sqrt{N}))$ be absolutely continuous with respect to $d\sigma^N$ with a probability density function F_N . We denote by

$$H_N(F_N) = H(F_N d\sigma^N | d\sigma^N) = H(\mu | \sigma^N).$$

Of special import to our work is the concept of marginals, and in particular first marginals.

Definition 1.3 Given $\mu \in P(\mathbb{R}^d)$ we define its k th marginal in the (i_1, \dots, i_k) th variables as the probability measure $\Pi_k^{(i_1, \dots, i_k)}(\mu)$ on \mathbb{R}^k satisfying

$$\Pi_k^{(i_1, \dots, i_k)}(\mu)(A_1 \times \dots \times A_k) = \mu(A^{(i_1, \dots, i_k)}), \tag{1.1}$$

where $A^{(i_1, \dots, i_k)} = \widetilde{A}_1 \times \dots \times \widetilde{A}_N$ with $\widetilde{A}_j = \begin{cases} A_l & j = i_l, l = 1, \dots, k \\ \mathbb{R} & j \neq i_1, \dots, i_k \end{cases}$.

It is important to note that even if a probability measure, μ , is supported on $\mathbb{S}^{N-1}(\sqrt{N})$, its k th marginal is well defined on \mathbb{R}^k whenever $k \leq N - 1$ and is supported in the ball of radius \sqrt{N} centered at the origin. Moreover, if μ is absolutely continuous with respect to the uniform probability measure $d\sigma^N$ then its k th marginal in the (i_1, \dots, i_k) th variables is absolutely continuous with respect to the Lebesgue measure on \mathbb{R}^k . We will denote by $\Pi_k^{(i_1, \dots, i_k)}(F_N)$ the probability density function of $\Pi_k^{(i_1, \dots, i_k)}(\mu)$.

From this point onward we will use the abusive notation of interchanging measures and their probability densities in all of our appropriate quantities; for instance, we will write $f \in P(\mathbb{R})$ when $f(v)dv \in P(\mathbb{R})$, etc.

We are now prepared to discuss the deviation from the equivalence of equilibrium principle, previously mentioned. It is simple to show (see the Appendix) that given $\mu \in P(\mathbb{R}^N)$ such that $d\mu = F_N dv$, with F_N having a finite second moment, one has that

$$\sum_{j=1}^N H\left(\Pi_1^{(j)}(F_N) \mid \gamma\right) \leq H(F_N \mid \gamma_N), \tag{1.2}$$

where $\gamma = \gamma_1$. This inequality is not too surprising and can be explained in physical terms: If one accepts that the entropy measures disorder in a given system, the right-hand side of the inequality represents the total disorder in the system of N objects that is described by F_N . The left-hand side, on the other hand, represents the sum of the *individual* disorders of the system. As by computing the latter we may have neglected some correlation terms between the objects of the system, the left-hand side needs to be less than the right-hand side. Moreover, equality can be attained only in the case where there are no correlations, i.e., the objects are independent. This indeed can be verified in the formal proof of the above.

Trying to generalize (1.2) one can define an appropriate first marginal on the sphere whenever F_N is a probability density function on $\mathbb{S}^{N-1}(\sqrt{N})$ by

$$F_j^{(N)}(v) = \int_{\mathbb{S}^{N-2}(\sqrt{N-v^2})} F_N d\sigma_{\sqrt{N-v^2}}^{N-1}, \tag{1.3}$$

where $d\sigma_r^k$ is the uniform probability measure on $\mathbb{S}^{k-1}(r)$. The expectation that (1.2) will be approximately true on the sphere was shown to be false in general. Indeed, in [3] the authors proved that

Theorem 1.4 *Let $F_N \in P(\mathbb{S}^{N-1}(\sqrt{N}))$. Then*

$$\sum_{i=1}^N \int_{\mathbb{S}^{N-1}(\sqrt{N})} F_j^{(N)} \log F_j^{(N)} d\sigma^N \leq 2H_N(F_N), \tag{1.4}$$

and the constant 2 is sharp.

The fact that the constant 2 is sharp shows that there is some fundamental difference between the sphere with uniform measure and \mathbb{R}^N with the Gaussian measure when one deals with all the objects of the statistical problem. An interesting problem to explore is identifying those quantities that are relevant to this deviation from the equivalence of ensembles idea.

Surprisingly enough, inequality (1.4) has ramifications in other fields in mathematics—in particular, kinetic theory and the study of Kac’s model. Kac’s model is a stochastic model of an average many-particle system that undergoes binary collisions from which a one-dimensional Boltzmann-like equation (called the Kac–Boltzmann equation) arises as a mean field limit. In his work, Kac had hoped to use his model, whose complexity comes from the number of particles and not any non-linearity, to solve unknown questions for the associated Boltzmann equation. Of particular interest to Kac was finding the rate of convergence to equilibrium. He suggested using the L^2 distance and the associated spectral gap of the evolution operator to tackle this particular problem. While the spectral gap was proved to be bounded from below uniformly in N (Kac’s conjecture), the L^2 distance was shown to be a catastrophic distance to consider under the setting of the model. A new distance, the relative entropy on the sphere, was investigated and with it the appropriate candidate for the rate of convergence: the entropy–entropy production ratio

$$\Gamma_N = \inf_{F_N} \frac{D_N(F_N)}{H_N(F_N)},$$

where $-D_N(F_N)$ is obtained by differentiating the entropy under Kac’s flow. For exponential decay of the entropy, one would hope to find a constant $C > 0$, independent of N , such that $\Gamma_N \geq C$. The existence of such a constant is known as the many-body Cercignani’s conjecture. Unfortunately, in [14] Villani has proven that

$$\Gamma_N \geq \frac{2}{N - 1}, \tag{1.5}$$

using a clever argument utilizing the heat semigroup on Kac’s sphere, and conjectured that $\Gamma_N = O\left(\frac{1}{N}\right)$, a claim that was essentially proved in [7]. Remarkably, Carlen showed in [2] that one can get (1.5) by using (1.4) and an inductive argument. The factor 2 plays a crucial role in the proof, and one notices that if it was replaced with $1 + \epsilon_N$, with ϵ_N converging to zero in a weak way, one would be able to prove Cercignani’s conjecture. This, as well as the investigation of the equivalence of ensembles, was the main motivation behind the investigation of the present work, and gave a framework from which the tools to prove our main theorems arose.

For more information about Kac’s model and the many-body Cercignani’s conjecture we refer the reader to [4, 8, 10, 11, 14].

We are finally ready to state the goal of the present work: Finding sufficient conditions on the probability density F_N on $\mathbb{S}^{N-1}\left(\sqrt{N}\right)$ under which (1.2) is indeed a good approximation to the appropriate spherical analogue in the sense that the constant 2 in (1.4) can be replaced by $1 + \epsilon_N$, with an explicit expression for ϵ_N .

Before we state our main theorems, we recall some notation from the theory of optimal transportation:

Definition 1.5 Let $\mu, \nu \in P(\mathbb{R}^d)$. The relative Fisher information of μ with respect to ν is defined as

$$I(\mu|\nu) = \begin{cases} \int_{\mathbb{R}^d} |\nabla \log h|^2 h d\nu & h = \frac{d\mu}{d\nu} \\ \infty & \text{otherwise.} \end{cases} \tag{1.6}$$

One can extend the definition of the relative Fisher information to $\mathbb{S}^{N-1}(\sqrt{N})$ in the case where $d\mu = F_N d\sigma^N$ and $d\nu = d\sigma^N$.

Definition 1.6 Let $F_N \in P(\mathbb{S}^{N-1}(\sqrt{N}))$. The Fisher information of F_N is defined as

$$I_N(F_N) = I_N(F_N d\sigma^N | d\sigma^N) = \int_{\mathbb{S}^{N-1}(\sqrt{N})} |\nabla_{\mathbb{S}} \log F_N|^2 F_N d\sigma^N, \tag{1.7}$$

where $\nabla_{\mathbb{S}}$ is the gradient on the sphere.

For more information about optimal transportation, its tools and applications we refer the reader to the excellent [12, 13].

Last, but not least, for any measurable, non-negative function f on \mathbb{R}^d we denote by

$$M_k(f) = \int_{\mathbb{R}^d} |v|^k f(v) d\nu \tag{1.8}$$

the k th moment of f .

The main theorems of this paper are:

Theorem 1.7 Let $F_N \in P(\mathbb{S}^{N-1}(\sqrt{N}))$ such that there exists $k > 2$ with

$$\mathcal{A}_k = \sup_N \frac{\sum_{j=1}^N M_k(\Pi_1^{(j)}(F_N))}{N} < \infty.$$

Assume in addition that

$$\mathcal{A}_I = \sup_N \frac{\sum_{i=1}^N I(\Pi_1^{(j)}(F_N))}{N} < \infty,$$

and that there exists $C_H > 0$ such that

$$\inf_N \frac{H_N(F_N)}{N} \geq C_H.$$

Then there exists an explicit ϵ_N , which goes to zero as a negative power of N and depends on $\mathcal{A}_k, \mathcal{A}_I, C_H$ and k , such that

$$\sum_{j=1}^N \int_{\mathbb{S}^{N-1}(\sqrt{N})} F_j^{(N)} \log F_j^{(N)} d\sigma^N \leq (1 + \epsilon_N) H_N(F_N). \tag{1.9}$$

Theorem 1.8 Let $F_N \in P(\mathbb{S}^{N-1}(\sqrt{N}))$ such that there exists $k > 2$ with

$$\mathcal{A}_k = \sup_N \frac{\sum_{j=1}^N M_k(\Pi_1^{(j)}(F_N))}{N} < \infty.$$

Assume in addition that there exists $2 < q < k$ such that

$$\mathcal{A}_q^P = \sup_N \frac{\sum_{j=1}^N P_q^{(j)}(F_N)}{N} < \infty,$$

where

$$P_q^{(j)}(F_N) = \int_{\mathbb{R}} \frac{\Pi_1^{(j)}(F_N)(v)}{\left(1 - \frac{v^2}{N}\right)^{\frac{q}{q-2}}} dv,$$

and that there exist constants $C_H, C_I > 0$ such that

$$\inf_N \frac{H_N(F_N)}{N} \geq C_H,$$

$$\sup_N \frac{I_N(F_N)}{N} \leq C_I.$$

Then there exists an explicit ϵ_N , which goes to zero as a negative power of N and depends on $\mathcal{A}_k, \mathcal{A}_q^P, C_I, C_H$ and k , such that

$$\sum_{j=1}^N \int_{\mathbb{S}^{N-1}(\sqrt{N})} F_j^{(N)} \log F_j^{(N)} d\sigma^N \leq (1 + \epsilon_N) H_N(F_N). \tag{1.10}$$

We'd like to point out a difference between our theorems: Theorem 1.7 requires an average bound on the Fisher information of the first marginals of F_N , a property that is not very intrinsic to the sphere. Theorem 1.8, on the other hand, relaxes this requirement and asks for information about the appropriate Fisher information on the sphere. However, as the gradient on the sphere of any function of one variable v_j is dampened near the poles $v_j = \pm\sqrt{N}$, an additional control condition near the poles is needed, which is where $P_q^{(j)}$ comes into play.

The idea of the proof of both theorems is to extend F_N from the sphere to \mathbb{R}^N where we are able to use (1.2). We shall call this extension *the Euclidean extension*. Once

that is done one investigates the connection between the marginals of the extension of F_N and F_N using an appropriate distance (the Wasserstein distance) and associate the entropies of the appropriate marginals using an HWI theorem. The final step involves finding the connection between the entropy of the marginal and the entropy of the marginal on the sphere.

The structure of the paper is as follows: In Sect. 2 we will describe the Euclidean extension, and see the connections between the first marginals and their moments, with respect to the original density. The entropic connection between the first marginals of the extension and the original density will be investigated in Sect. 3, while the entropic connection between the first marginals and the first marginals on the sphere will be shown in Sect. 4. We will prove our main theorems in Sect. 5 and give a non-trivial example for when the conditions of the theorems are satisfied in Sect. 6. We then conclude the paper with a few final remarks in Sect. 7 and deal with a few technical computations in the Appendix.

2 The Euclidean Extension and Marginal Relation

The first step on the path to improve (1.4) is passing from the sphere to Euclidean space. This is done by extending a given $F_N \in P\left(\mathbb{S}^{N-1}\left(\sqrt{N}\right)\right)$ to a function on \mathbb{R}^N , \widetilde{F}_N , in a way that is compatible with the entropy.

Definition 2.1 Given $F_N \in P\left(\mathbb{S}^{N-1}\left(\sqrt{N}\right)\right)$, its Euclidean extension \widetilde{F}_N is defined as

$$\widetilde{F}_N(v) = F_N\left(\sqrt{N}\frac{v}{|v|}\right) \cdot \gamma_N(v), \tag{2.1}$$

with $v \in \mathbb{R}^N \setminus \{0\}$.

Lemma 2.2 $\widetilde{F}_N \in P\left(\mathbb{R}^N\right)$ and

$$H\left(\widetilde{F}_N|\gamma_N\right) = H_N\left(F_N\right).$$

Proof Using spherical coordinates, the fact that $F_N\left(\sqrt{N}\frac{v}{|v|}\right)$ depends only on the angular variable and the fact that γ_N is radial we see that:

$$\begin{aligned} H\left(\widetilde{F}_N|\gamma_N\right) &= \int_{\mathbb{R}^N} F_N\left(\sqrt{N}\frac{v}{|v|}\right) \log\left(F_N\left(\sqrt{N}\frac{v}{|v|}\right)\right) \gamma_N(v) dv \\ &= \left(\int_{\mathbb{S}^{N-1}} F_N\left(\sqrt{N}\frac{v}{|v|}\right) \log\left(F_N\left(\sqrt{N}\frac{v}{|v|}\right)\right) d\sigma_1^N\right) \\ &\quad \left(|\mathbb{S}^{N-1}| \int_0^\infty \frac{r^{N-1}}{(2\pi)^{\frac{N}{2}}} e^{-\frac{r^2}{2}} dr\right) \\ &= H_N\left(F_N\right), \end{aligned}$$

since

$$1 = \int_{\mathbb{R}^N} \gamma_N(v)dv = |\mathbb{S}^{N-1}| \int_0^\infty \frac{r^{N-1}}{(2\pi)^{\frac{N}{2}}} e^{-\frac{r^2}{2}} dr. \tag{2.2}$$

Using the same argument one can easily show that \widetilde{F}_N is indeed a probability density. \square

Now that we have a possible extension at hand, the next step we'd like to explore is the relation between its first marginals and those of the original function. We start by recalling the following simple Fubini–Tonelli type theorem on the sphere (see [7], for instance):

$$\int_{\mathbb{S}^{N-1}(r)} F_N d\sigma_r^N = \frac{|\mathbb{S}^{N-k-1}|}{|\mathbb{S}^{N-1}|} \frac{1}{r^{N-2}} \int_{\sum_{i=1}^k v_i^2 \leq r^2} \left(r^2 - \sum_{i=1}^k v_i^2 \right)^{\frac{N-k-2}{2}} \left(\int_{\mathbb{S}^{N-k-1}(\sqrt{r^2 - \sum_{i=1}^k v_i^2})} F_N d\sigma^{\frac{N-k}{2}} \right) dv_1 \dots dv_k. \tag{2.3}$$

Formula (2.3) allows us to write a concrete expression for the k th marginal of a probability density function $F_N \in P(\mathbb{S}^{N-1}(\sqrt{N}))$ in its (i_1, \dots, i_k) variables whenever $k \leq N - 1$. Indeed, one easily sees that

$$\Pi_k^{(i_1, \dots, i_k)}(F_N)(v_{i_1}, \dots, v_{i_k}) = \frac{|\mathbb{S}^{N-k-1}|}{|\mathbb{S}^{N-1}|} \frac{1}{N^{\frac{k}{2}}} \left(1 - \frac{\sum_{l=1}^k v_{i_l}^2}{N} \right)_+^{\frac{N-k-2}{2}} \left(\int_{\mathbb{S}^{N-k-1}(\sqrt{N - \sum_{i=1}^k v_{i_l}^2})} F_N d\sigma^{\frac{N-k}{2}} \right). \tag{2.4}$$

Using this, we can conclude the following:

Lemma 2.3 *Let $F_N \in P(\mathbb{S}^{N-1}(\sqrt{N}))$. Then, the k th marginal of \widetilde{F}_N in the (i_1, \dots, i_k) variables is given by*

$$\begin{aligned} & \Pi_k^{(i_1, \dots, i_k)}(\widetilde{F}_N)(v_1, \dots, v_k) \\ &= |\mathbb{S}^{N-1}| N^{\frac{k}{2}} \int_0^\infty r \frac{\left(\sum_{l=1}^k v_{i_l}^2 + r^2 \right)^{\frac{N-k-2}{2}}}{(2\pi)^{\frac{N}{2}}} e^{-\frac{r^2 + \sum_{l=1}^k v_{i_l}^2}{2}} \\ & \Pi_k^{(i_1, \dots, i_k)}(F_N) \left(\frac{\sqrt{N}v_{i_1}}{\sqrt{\sum_{l=1}^k v_{i_l}^2 + r^2}}, \dots, \frac{\sqrt{N}v_{i_k}}{\sqrt{\sum_{l=1}^k v_{i_l}^2 + r^2}} \right) dr. \end{aligned} \tag{2.5}$$

Proof By its definition

$$\Pi_k^{(i_1, \dots, i_k)}(\tilde{F}_N)(v_{i_1}, \dots, v_{i_k}) = \int_{\mathbb{R}^{N-k}} F_N\left(\sqrt{N} \frac{v}{|v|}\right) \gamma_N(v_1, \dots, v_N) d\tilde{v}_{i_1, \dots, v_{i_k}}$$

where $d\tilde{v}_{i_1, \dots, v_{i_k}}$ represents dv excluding $dv_{i_1} \dots dv_{i_k}$. For the sake of simplicity of notation we'll assume that $i_l = l$. We find that

$$\begin{aligned} & \Pi_k^{(1, \dots, k)}(\tilde{F}_N)(v_1, \dots, v_k) \\ &= \int_{\mathbb{R}^{N-k}} F_N\left(\frac{\sqrt{N}v_1}{\sqrt{\sum_{i=1}^k v_i^2 + \sum_{i=k+1}^N v_i^2}}, \dots, \frac{\sqrt{N}v_N}{\sqrt{\sum_{i=1}^k v_i^2 + \sum_{i=k+1}^N v_i^2}}\right) \\ & \quad \frac{e^{-\frac{\sum_{i=1}^k v_i^2 + \sum_{i=k+1}^N v_i^2}{2}}}{(2\pi)^{\frac{N}{2}}} dv_{k+1} \dots dv_N \\ &= |\mathbb{S}^{N-k-1}| \int_{[0, \infty)} dr \frac{r^{N-k-1} e^{-\frac{\sum_{i=1}^k v_i^2 + r^2}{2}}}{(2\pi)^{\frac{N}{2}}} \\ & \quad \left(\int_{\mathbb{S}^{N-k-1}} \left(\frac{\sqrt{Nr}}{\sqrt{\sum_{i=1}^k v_i^2 + r^2}} \right) F_N\left(\frac{\sqrt{N}v_1}{\sqrt{\sum_{i=1}^k v_i^2 + r^2}}, \dots, \frac{\sqrt{N}v_k}{\sqrt{\sum_{i=1}^k v_i^2 + r^2}}, \sigma\right) \right. \\ & \quad \left. d\sigma^{N-k} \left(\frac{\sqrt{Nr}}{\sqrt{\sum_{i=1}^k v_i^2 + r^2}} \right) \right). \end{aligned}$$

Using (2.2) and (2.4) we find that the above equals

$$\begin{aligned} & |\mathbb{S}^{N-1}| N^{\frac{k}{2}} \\ & \int_{[0, \infty)} dr \frac{r^{N-k-1} e^{-\frac{\sum_{i=1}^k v_i^2 + r^2}{2}}}{(2\pi)^{\frac{N}{2}}} \frac{\Pi_k^{(1, \dots, k)}(F_N)\left(\frac{\sqrt{N}v_1}{\sqrt{\sum_{i=1}^k v_i^2 + r^2}}, \dots, \frac{\sqrt{N}v_k}{\sqrt{\sum_{i=1}^k v_i^2 + r^2}}\right)}{\left(1 - \frac{\sum_{i=1}^k v_i^2}{\sum_{i=1}^k v_i^2 + r^2}\right)_+^{\frac{N-k-2}{2}}}, \end{aligned}$$

from which the result follows. □

Of particular interest is the case of the first marginal in the j th variable, $\Pi_1^{(j)}(\tilde{F}_N)$. Using Lemma 2.3 we obtain

Corollary 2.4 *Let $F_N \in P(\mathbb{S}^{N-1}(\sqrt{N}))$. Then*

$$\Pi_1^{(j)}(\tilde{F}_N)(v) = \frac{|\mathbb{S}^{N-1}| N^{\frac{N}{2}}}{(2\pi)^{\frac{N}{2}}} \int_0^{\text{sgn}(v)\sqrt{N}} \Pi_1^{(j)}(F_N)(x) \frac{v^{N-1}}{x^N} e^{-\frac{Nv^2}{2x^2}} dx. \tag{2.6}$$

Proof From (2.5) we know that

$$\Pi_1^{(j)}(\tilde{F}_N)(v) = |\mathbb{S}^{N-1}| \sqrt{N} \int_0^\infty \Pi_1^{(j)}(F_N)\left(\frac{\sqrt{N}v}{\sqrt{v^2+r^2}}\right) \frac{r(v^2+r^2)^{\frac{N-3}{2}}}{(2\pi)^{\frac{N}{2}}} e^{-\frac{r^2+v^2}{2}} dr.$$

Using the change of variables $x = \frac{\sqrt{N}v}{\sqrt{v^2+r^2}}$ we find that

$$\begin{aligned} \Pi_1^{(j)}(\tilde{F}_N)(v) &= |\mathbb{S}^{N-1}| \sqrt{N} \int_0^{\text{sgn}(v)\sqrt{N}} \Pi_1^{(j)}(F_N)(x) \\ &\quad \left(\frac{Nv^2}{x^2}\right)^{\frac{N-3}{2}} \frac{Nv^2}{x^3 (2\pi)^{\frac{N}{2}}} e^{-\frac{Nv^2}{2x^2}} dx \\ &= \frac{|\mathbb{S}^{N-1}| N^{\frac{N}{2}}}{(2\pi)^{\frac{N}{2}}} \int_0^{\text{sgn}(v)\sqrt{N}} \Pi_1^{(j)}(F_N)(x) \frac{v^{N-1}}{x^N} e^{-\frac{Nv^2}{2x^2}} dx, \end{aligned}$$

completing the proof. □

An interesting application of Corollary 2.4 is a moment connection between $\Pi_1^{(j)}(F_N)$ and $\Pi_1^{(j)}(\tilde{F}_N)$.

Lemma 2.5 *Let $F_N \in P(\mathbb{S}^{N-1}(\sqrt{N}))$. Then*

$$\int_{\mathbb{R}} |v|^m \Pi_1^{(j)}(\tilde{F}_N)(v) dv = \left(\frac{2}{N}\right)^{\frac{m}{2}} \frac{\Gamma\left(\frac{N+m}{2}\right)}{\Gamma\left(\frac{N}{2}\right)} \int_{-\sqrt{N}}^{\sqrt{N}} |v|^m \Pi_1^{(j)}(F_N)(v) dv. \tag{2.7}$$

Proof Using (2.6), we have that

$$\begin{aligned} &\int_{\mathbb{R}} |v|^m \Pi_1(\tilde{F}_N)(v) dv \\ &= \frac{|\mathbb{S}^{N-1}| N^{\frac{N}{2}}}{(2\pi)^{\frac{N}{2}}} \left(\int_0^\infty \int_0^{\sqrt{N}} \Pi_1(F_N)(x) \frac{v^{N+m-1}}{x^N} e^{-\frac{Nv^2}{2x^2}} dx dv \right. \\ &\quad \left. - \int_{-\infty}^0 \int_{-\sqrt{N}}^0 \Pi_1(F_N)(x) \frac{(-1)^m v^{N+m-1}}{x^N} e^{-\frac{Nv^2}{2x^2}} dx dv \right) \\ &= \frac{|\mathbb{S}^{N-1}| (2\pi)^{\frac{m}{2}}}{N^{\frac{m}{2}}} \int_{-\sqrt{N}}^{\sqrt{N}} |x|^m \Pi_1(F_N)(x) \left(\frac{1}{(2\pi)^{\frac{N+m}{2}}} \int_0^\infty y^{N+m-1} e^{-\frac{y^2}{2}} dy \right) dx \end{aligned}$$

$$= \frac{|\mathbb{S}^{N-1}| (2\pi)^{\frac{m}{2}}}{N^{\frac{m}{2}} |\mathbb{S}^{N+m-1}|} \int_{-\sqrt{N}}^{\sqrt{N}} |x|^m \Pi_1(F_N)(x) dx,$$

where we have used (2.2). The result follows from the formula

$$|\mathbb{S}^{N-1}| = \frac{2\pi^{\frac{N}{2}}}{\Gamma(\frac{N}{2})}. \tag{2.8}$$

□

Lemma 2.5 implies the following:

Corollary 2.6 *For any $k > 0$ there exists $C_k > 0$, independent of N , such that*

$$M_k\left(\Pi_1^{(j)}(\widetilde{F}_N)\right) \leq C_k M_k\left(\Pi_1^{(j)}(F_N)\right). \tag{2.9}$$

Moreover, when $k = 2$ there is equality in (2.9) with $C_2 = 1$.

Proof From (2.7) we see that choosing

$$C_k = \sup_N \left(\frac{2}{N}\right)^{\frac{k}{2}} \frac{\Gamma(\frac{N+k}{2})}{\Gamma(\frac{N}{2})}$$

proves the claim. The boundedness of $\left(\frac{2}{N}\right)^{\frac{k}{2}} \frac{\Gamma(\frac{N+k}{2})}{\Gamma(\frac{N}{2})}$ is clear from the asymptotic expansion of the Gamma function

$$\Gamma(z) = z^{z-\frac{1}{2}} e^{-z} \sqrt{2\pi} \left(1 + \frac{1}{12z} + \dots\right). \tag{2.10}$$

If $k = 2l$ then

$$\Gamma\left(\frac{N+k}{2}\right) = \Gamma\left(\frac{N}{2} + l\right) = \left(\prod_{i=0}^{l-1} \left(\frac{N}{2} - i\right)\right) \Gamma\left(\frac{N}{2}\right).$$

In this case,

$$M_k\left(\Pi_1^{(j)}(\widetilde{F}_N)\right) = \left(\prod_{i=0}^{\frac{k}{2}-1} \left(1 - \frac{2i}{N}\right)\right) M_k\left(\Pi_1^{(j)}(F_N)\right) \leq M_k\left(\Pi_1^{(j)}(F_N)\right),$$

with equality if and only if $k = 2$.

□

3 The Entropy Relation: From Marginals to the Marginals of the Extension

Now that we have managed to extend our probability density from $\mathbb{S}^{N-1}(\sqrt{N})$ to \mathbb{R}^N we would like to find out how much information we may have “lost” during that process, at least in the sense of the entropy functional. The main theoretical tool to connect between the two will be the HWI inequality (see [6, 12, 13]). In this section we will slowly investigate the quantities that will play a role in the final connection between the entropies, namely the Wasserstein distance and the Fisher information, and eventually quantify the “loss” in the transition followed by our extension.

We start by reminding the reader of the definition of the Wasserstein distance:

Definition 3.1 Let X be a Polish space with a metric d and let μ, ν be two probability measures on X . For any $q \geq 1$ the Wasserstein distance of order q between μ and ν is defined as

$$W_q(\mu, \nu) = \left(\inf_{\pi \in \Pi(\mu, \nu)} \int_{X \times X} d^q(x, y) d\pi(x, y) \right)^{\frac{1}{q}}, \tag{3.1}$$

where $\Pi(\mu, \nu)$, the space of coupling, is the space of all probability measures on $X \times X$ with marginals μ and ν respectively.

Our first lemma shows that the Wasserstein distance of order 1 is indeed the right distance to measure the difference between the marginal of F_N and that of its extension:

Lemma 3.2 Let $F_N \in P(\mathbb{S}^{N-1}(\sqrt{N}))$. Then for any $j = 1, \dots, N$

$$W_1\left(\Pi_1^{(j)}(F_N), \Pi_1^{(j)}(\widetilde{F}_N)\right) \leq \frac{M_2\left(\Pi_1^{(j)}(F_N)\right)^{\frac{1}{2}}}{\sqrt{2N}} (1 + \tau_N), \tag{3.2}$$

where $\tau_N \xrightarrow{N \rightarrow \infty} 0$ as N goes to infinity, is given explicitly and independently of F_N .

Proof The proof relies on the famous Kantorovich–Rubinstein formula (see [12]): For any $\mu, \nu \in P(X)$, where X is a Polish space,

$$W_1(\mu, \nu) = \sup \left(\int_X \psi(x) (d\mu - d\nu)(x) \right),$$

where the supremum is taken over all 1-Lipschitz functions ψ .

For any $\phi \in C_b(\mathbb{R})$ we find that

$$\begin{aligned} \int_{\mathbb{R}} \phi(v) \Pi_1^{(j)}(\widetilde{F}_N) dv &= \frac{|\mathbb{S}^{N-1}|}{(2\pi)^{\frac{N}{2}}} N^{\frac{N}{2}} \int_0^\infty \\ &\int_0^{\sqrt{N}} \phi(v) \Pi_1^{(j)}(F_N)(x) \frac{v^{N-1}}{x^N} e^{-\frac{Nv^2}{2x^2}} dv dx \\ &+ \frac{|\mathbb{S}^{N-1}|}{(2\pi)^{\frac{N}{2}}} N^{\frac{N}{2}} \int_{-\infty}^0 \int_{-\sqrt{N}}^0 \phi(v) \Pi_1^{(j)}(F_N)(x) \left(-\frac{v^{N-1}}{x^N}\right) e^{-\frac{Nv^2}{2x^2}} dv dx \\ &= \frac{|\mathbb{S}^{N-1}|}{(2\pi)^{\frac{N}{2}}} \int_0^\infty \int_{-\sqrt{N}}^{\sqrt{N}} \phi\left(\frac{yx}{\sqrt{N}}\right) \Pi_1^{(j)}(F_N)(x) y^{N-1} e^{-\frac{y^2}{2}} dy dx. \end{aligned}$$

Using (2.2) and the fact that $\Pi_1^{(j)}(F_N)$ is supported in $[-\sqrt{N}, \sqrt{N}]$ we see that

$$\begin{aligned} &\left| \int_{\mathbb{R}} \phi(x) \Pi_1^{(j)}(F_N)(x) dx - \int_{\mathbb{R}} \phi(v) \Pi_1^{(j)}(\widetilde{F}_N)(v) dv \right| \\ &\leq \frac{|\mathbb{S}^{N-1}|}{(2\pi)^{\frac{N}{2}}} \int_0^\infty \int_{-\sqrt{N}}^{\sqrt{N}} \left| \phi\left(\frac{yx}{\sqrt{N}}\right) - \phi(x) \right| \Pi_1^{(j)}(F_N)(x) y^{N-1} e^{-\frac{y^2}{2}} dy dx. \end{aligned}$$

If in addition ϕ is 1-Lipschitz we have that

$$\begin{aligned} &\left| \int_{\mathbb{R}} \phi(x) \Pi_1^{(j)}(F_N)(x) dx - \int_{\mathbb{R}} \phi(v) \Pi_1^{(j)}(\widetilde{F}_N)(v) dv \right| \\ &\leq \left(\int_{-\sqrt{N}}^{\sqrt{N}} |x| \Pi_1^{(j)}(F_N)(x) dx \right) \left(\frac{|\mathbb{S}^{N-1}|}{(2\pi)^{\frac{N}{2}}} \int_0^\infty \left| \frac{y}{\sqrt{N}} - 1 \right| y^{N-1} e^{-\frac{y^2}{2}} dy dx \right) \\ &\leq M_2 \left(\Pi_1^{(j)}(F_N) \right)^{\frac{1}{2}} \left(\frac{|\mathbb{S}^{N-1}|}{(2\pi)^{\frac{N}{2}}} \int_0^\infty \left(\frac{y}{\sqrt{N}} - 1 \right)^2 y^{N-1} e^{-\frac{y^2}{2}} dy \right)^{\frac{1}{2}}. \end{aligned}$$

It is easy to see that

$$\int_0^\infty \left(\frac{y}{\sqrt{N}} - 1 \right)^2 y^{N-1} e^{-\frac{y^2}{2}} dy = 2 \frac{(2\pi)^{\frac{N}{2}}}{|\mathbb{S}^{N-1}|} \left(1 - \frac{\sqrt{2\pi}}{\sqrt{N}} \frac{|\mathbb{S}^{N-1}|}{|\mathbb{S}^N|} \right).$$

The result follows from a simple asymptotic expansion of $|\mathbb{S}^{N-1}|$ using (2.8) and (2.10). □

While the Wasserstein distance of order 1 was one that yielded the desired closeness between the appropriate marginals, it is not one that is suited to associate with the entropy via the so-called HWI inequality. In order to use the aforementioned inequality we will require higher orders of Wasserstein distances. Our next lemma, which is a simple extension of a result proved by Hauray and Mischler in [10], allows us to make

the connection between W_1 and $W_q, q \geq 1$, as long as we have additional moment control.

Lemma 3.3 *Let $f, g \in P(\mathbb{R})$ and let $k > 0$. Denote by*

$$\mathcal{M}_k = \mathcal{M}_k(f) + \mathcal{M}_k(g) = \int_{\mathbb{R}} (1 + |v|^2)^{\frac{k}{2}} f(v)dv + \int_{\mathbb{R}} (1 + |v|^2)^{\frac{k}{2}} g(v)dv.$$

Then, for any $k > q \geq 2$ one has that

$$W_q(f, g) \leq 2^{1+\frac{1}{q}} \mathcal{M}_k^{\frac{1}{k}} W_1(f, g)^{\frac{1}{q}-\frac{1}{k}}. \tag{3.3}$$

Proof Denote by $d(x, y) = \min(|x - y|, 1)$ and by \widetilde{W}_1 the Wasserstein distance of order 1 associated with d . We claim that for all $q \geq 1$ and $R \geq 1$

$$|x - y|^q \leq R^q d(x, y) + \frac{2^k}{R^{k-q}} (|x|^k + |y|^k), \tag{3.4}$$

and leave the proof of this inequality to the Appendix. Integrating (3.4) against any $\pi \in \Pi(\mu, \nu)$ gives us

$$W_q^q(\mu, \nu) \leq R^q \widetilde{W}_1(\mu, \nu) + \frac{2^k}{R^{k-q}} \mathcal{M}_k(\mu, \nu).$$

The choice $R = 2 \frac{\mathcal{M}_k^{\frac{1}{k}}(\mu, \nu)}{\widetilde{W}_1^{\frac{1}{k}}(\mu, \nu)} \geq 1$ yields

$$W_q(\mu, \nu) \leq 2(2)^{\frac{1}{q}} \mathcal{M}_k^{\frac{1}{k}}(\mu, \nu) \widetilde{W}_1^{\frac{1}{q}-\frac{1}{k}}(\mu, \nu),$$

from which the result follows as $\widetilde{W}_1 \leq W_1$. □

Corollary 3.4 *Let $F_N \in P(\mathbb{S}^{N-1}(\sqrt{N}))$ be such that $M_k(\Pi_1^{(j)}(F_N)) < \infty$ for some $k > 2$. Then for any $2 \leq q < k$*

$$W_q(\Pi_1^{(j)}(F_N), \Pi_1^{(j)}(\widetilde{F}_N)) \leq 2^{\frac{3}{2}+\frac{1}{q}} \left(1 + \frac{C_k}{2}\right)^{\frac{1}{k}} \tag{3.5}$$

$$\left(1 + M_k(\Pi_1^{(j)}(F_N))\right)^{\frac{1}{k}} \frac{M_2(\Pi_1^{(j)}(F_N))^{\frac{1}{2q}-\frac{1}{2k}}}{(2N)^{\frac{1}{2q}-\frac{1}{2k}}} \left(1 + O\left(\frac{1}{N}\right)\right),$$

with $C_k = \sup_N \left(\frac{2}{N}\right)^{\frac{k}{2}} \frac{\Gamma(\frac{N+k}{2})}{\Gamma(\frac{N}{2})}$.

Proof Using the notation of Lemma 3.3 we have that

$$\begin{aligned} \mathcal{M}_k &\leq 2^{\frac{k}{2}-1} \left(2 + M_k \left(\Pi_1^{(j)}(F_N) \right) + M_k \left(\Pi_1^{(j)}(\widetilde{F}_N) \right) \right) \\ &\leq 2^{\frac{k}{2}} \left(1 + \frac{C_k}{2} \right) \left(1 + M_k \left(\Pi_1^{(j)}(F_N) \right) \right). \end{aligned}$$

Combining this with Lemmas 3.2 and 3.3 yields the desired result. □

The next ingredient of the proof that we need is the Fisher information. While the “normal” one, defined in Definition 1.5 and used in Theorem 1.7 requires no further discussion, we will require the following lemmas to deal with the Fisher information on the sphere.

Lemma 3.5 *Let $F_N \in P \left(\mathbb{S}^{N-1} \left(\sqrt{N} \right) \right)$. Then*

$$\begin{aligned} \Pi_1^{(j)}(F_N)(v) &= \frac{|\mathbb{S}^{N-2}|}{|\mathbb{S}^{N-1}| \sqrt{N}} \left(1 - \frac{v^2}{N} \right)_+^{\frac{N-3}{2}} F_j^{(N)}(v) \\ &= \frac{1}{\sqrt{2\pi}} \left(1 - \frac{3}{4N} + o\left(\frac{1}{N}\right) \right) \left(1 - \frac{v^2}{N} \right)_+^{\frac{N-3}{2}} F_j^{(N)}(v). \end{aligned} \tag{3.6}$$

Proof Equality (3.6) follows immediately from (2.4) with $k = 1$ and an asymptotic expansion of $|\mathbb{S}^{N-1}|$. □

Lemma 3.6 *Let $F_N \in P \left(\mathbb{S}^{N-1} \left(\sqrt{N} \right) \right)$ such that $I_N \left(F_j^{(N)} \right) < \infty$. Then*

$$\begin{aligned} I_N \left(F_j^{(N)} \right) &= \int_{\mathbb{R}} \left(1 - \frac{v^2}{N} \right) \left| \frac{d}{dv} \log \left(\Pi_1^{(j)}(F_N)(v) \right) \right|^2 \Pi_1^{(j)}(F_N)(v) \\ &\quad - 2 \frac{N-3}{N} + \left(\frac{N-3}{N} \right)^2 \int_{\mathbb{R}} \frac{v^2 \Pi_1^{(j)}(F_N)(v)}{\left(1 - \frac{v^2}{N} \right)} dv. \end{aligned} \tag{3.7}$$

Proof Denote by $L_{i,j} = \frac{1}{\sqrt{N}} (v_i \partial_j - v_j \partial_i)$. For any F on $\mathbb{S}^{N-1} \left(\sqrt{N} \right)$ we have that

$$\begin{aligned} I_N(F) &= \int_{\mathbb{S}^{N-1}(\sqrt{N})} \frac{|\nabla_{\mathbb{S}} F|^2}{F} d\sigma^N = \frac{1}{2} \sum_{i \neq j} \int_{\mathbb{S}^{N-1}(\sqrt{N})} \frac{|L_{i,j} F|^2}{F} d\sigma^N \\ &= \frac{1}{2} \sum_{i \neq j} \int_{\mathbb{S}^{N-1}(\sqrt{N})} |L_{i,j} \log F|^2 F d\sigma^N. \end{aligned}$$

If $F = f_j$, a function depending only on v_j , we find that

$$\sum_{i \neq k} |L_{i,k} f_j|^2 = \frac{2}{N} \sum_{i \neq j} v_i^2 \left(\frac{d}{dv_j} f_j \right)^2 = 2 \left(1 - \frac{v_j^2}{N} \right) \left(\frac{d}{dv_j} f_j \right)^2.$$

Thus, using (3.6) we find that

$$\begin{aligned}
 I_N \left(F_j^{(N)} \right) &= \int_{\mathbb{S}^{N-1}(\sqrt{N})} \left(1 - \frac{v^2}{N} \right) \left| \frac{d}{dv} \log \left(\frac{\Pi_1^{(j)}(F_N)(v)}{\frac{|\mathbb{S}^{N-2}|}{|\mathbb{S}^{N-1}|\sqrt{N}} \left(1 - \frac{v^2}{N} \right)^{\frac{N-3}{2}}} \right) \right|^2 F_j^{(N)}(v) d\sigma^N \\
 &= \int_{\mathbb{R}} \left(1 - \frac{v^2}{N} \right) \left| \frac{d}{dv} \log \left(\Pi_1^{(j)}(F_N)(v) \right) + \frac{(N-3)v}{N \left(1 - \frac{v^2}{N} \right)} \right|^2 \Pi_1^{(j)}(F_N)(v) dv \\
 &= \int_{\mathbb{R}} \left(1 - \frac{v^2}{N} \right) \left| \frac{d}{dv} \log \left(\Pi_1^{(j)}(F_N)(v) \right) \right|^2 \Pi_1^{(j)}(F_N)(v) dv \\
 &\quad + 2 \frac{N-3}{N} \int_{\mathbb{R}} v \frac{d}{dv} \left(\Pi_1^{(j)}(F_N)(v) \right) dv \\
 &\quad + \left(\frac{N-3}{N} \right)^2 \int_{\mathbb{R}} \frac{v^2 \Pi_1^{(j)}(F_N)(v)}{1 - \frac{v^2}{N}} dv,
 \end{aligned}$$

where we have used (2.3) in the second line. The result follows by a simple integration by parts. □

Using our acquired knowledge until this point we can now find a quantitative estimation in the difference of the entropies of the marginals and the marginals of the extension.

Theorem 3.7 *Let $F_N \in P \left(\mathbb{S}^{N-1} \left(\sqrt{N} \right) \right)$ such that $M_k \left(\Pi_1^{(j)}(F_N) \right) < \infty$ for some $k > 2$.*

(i) If $I \left(\Pi_1^{(j)}(F_N) \right) < \infty$ then there exists $C_2 > 0$, independent of N and F_N , such that

$$\begin{aligned}
 H \left(\Pi_1^{(j)}(F_N) | \gamma \right) &\leq H \left(\Pi_1^{(j)}(\widetilde{F}_N) | \gamma \right) \\
 &\quad + 4C_2 \left(1 + \frac{C_k}{2} \right)^{\frac{1}{k}} \left(1 + M_k \left(\Pi_1^{(j)}(F_N) \right) \right)^{\frac{1}{k}} \\
 &\quad \frac{\left(I \left(\Pi_1^{(j)}(F_N) \right) + M_2 \left(\Pi_1^{(j)}(F_N) \right) - 2 \right)^{\frac{1}{2}} M_2 \left(\Pi_1^{(j)}(F_N) \right)^{\frac{1}{4} - \frac{1}{2k}}}{(2N)^{\frac{1}{4} - \frac{1}{2k}}} \\
 &= H \left(\Pi_1^{(j)}(\widetilde{F}_N) | \gamma \right) + \epsilon_N^{(1,j)},
 \end{aligned} \tag{3.8}$$

where $C_k = \sup_N \left(\frac{2}{N} \right)^{\frac{k}{2}} \frac{\Gamma \left(\frac{N+k}{2} \right)}{\Gamma \left(\frac{N}{2} \right)}$.

(ii) If $I_N \left(\Pi_1^{(j)} (F_N) \right) < \infty$ and there exists $2 < q < k$ such that

$$P_q^{(j)} (F_N) = \int_{\mathbb{R}} \frac{\Pi_1^{(j)} (F_N) (v)}{\left(1 - \frac{v^2}{N} \right)^{\frac{q}{q-2}}} dv < \infty \tag{3.9}$$

then there exists $C_2 > 0$, independent of N and F_N , such that

$$\begin{aligned} H \left(\Pi_1^{(j)} (F_N) | \gamma \right) &\leq H \left(\Pi_1^{(j)} (\widetilde{F}_N) | \gamma \right) + 2^{\frac{3}{2} + \frac{2}{q}} C_2 \left(1 + \frac{C_k}{2} \right)^{\frac{1}{k}} \\ &\left(\left(I_N (F_j^{(N)}) + 2 \right)^{\frac{q}{2(q-1)}} \left(P_q^{(j)} \right)^{\frac{q-2}{2(q-1)}} + 1 + M_2 \left(\Pi_1^{(j)} (F_N) \right) \right)^{\frac{q-1}{q}} \\ &\left(1 + M_k \left(\Pi_1^{(j)} (F_N) \right) \right)^{\frac{1}{k}} \frac{M_2 \left(\Pi_1^{(j)} (F_N) \right)^{\frac{1}{2q} - \frac{1}{2k}}}{(2N)^{\frac{1}{2q} - \frac{1}{2k}}} \\ &= H \left(\Pi_1^{(j)} (\widetilde{F}_N) | \gamma \right) + \epsilon_N^{(2,j)}, \end{aligned} \tag{3.10}$$

where $C_k = \sup_N \left(\frac{2}{N} \right)^{\frac{k}{2}} \frac{\Gamma \left(\frac{N+k}{2} \right)}{\Gamma \left(\frac{N}{2} \right)}$.

Proof (i) The HWI inequality states that

$$H(f | \gamma) \leq H(g | \gamma) + \sqrt{I(f | \gamma)} W_2(f, g).$$

Together with the simple identity for $f \in P(\mathbb{R})$

$$I(f | \gamma) = I(f) + \int_{\mathbb{R}} v^2 f(v) dv - 2,$$

Corollary 3.4 with $q = 2$, and the fact that the $O\left(\frac{1}{N}\right)$ term in (3.5) was independent in F_N , we conclude the result.

(ii) This part of the theorem requires a slight modification of the HWI inequality. Following the proof of the inequality (see, for instance, [6, 13]), one notices that replacing the Cauchy–Schwarz inequality with the Hölder inequality (and using the uniqueness of the transportation map if needed) gives us that for any $1 < p < \infty$

$$H(f | \gamma) \leq H(g | \gamma) + \left(\int_{\mathbb{R}} \left| \frac{d}{dv} \log \left(\frac{f(v)}{\gamma(v)} \right) \right|^p f(v) dv \right)^{\frac{1}{p}} W_q(f, g),$$

where q is the Hölder conjugate of p . For $1 \leq p < 2$ we find that

$$\begin{aligned} & \int_{\mathbb{R}} \left| \frac{d}{dv} \log \left(\frac{f(v)}{\gamma(v)} \right) \right|^p f(v) dv \\ &= \int_{\mathbb{R}} \left| \frac{d}{dv} \log f(v) + v \right|^p f(v) dv \\ &\leq 2^{p-1} \left(\int_{\mathbb{R}} \left| \frac{d}{dv} \log f(v) \right|^p f(v) dv + \int_{\mathbb{R}} |v|^p f(v) dv \right) \\ &\leq 2^{p-1} \left(\int_{\mathbb{R}} \left| \frac{d}{dv} \log f(v) \right|^p f(v) dv + 1 + M_2(f) \right), \end{aligned}$$

and if in addition $f \in P(\mathbb{R})$ is supported in $[-\sqrt{N}, \sqrt{N}]$ then

$$\begin{aligned} & \int_{\mathbb{R}} \left| \frac{d}{dv} \log f(v) \right|^p f(v) dv \\ &\leq \left(\int_{\mathbb{R}} \left(1 - \frac{v^2}{N} \right) \left| \frac{d}{dv} \log f(v) \right|^2 f(v) dv \right)^{\frac{p}{2}} \left(\int_{\mathbb{R}} \frac{f(v)}{\left(1 - \frac{v^2}{N} \right)^{\frac{p}{2-p}}} dv \right)^{\frac{2-p}{2}}. \end{aligned}$$

We conclude that for $p = \frac{q}{q-1}$, where q is as in (3.9), one has that

$$\begin{aligned} & H \left(\Pi_1^{(j)}(F_N) \mid \gamma \right) \leq H \left(\Pi_1^{(j)}(\widetilde{F}_N) \mid \gamma \right) \\ &+ 2^{\frac{1}{q}} \left(\left(\int_{\mathbb{R}} \left(1 - \frac{v^2}{N} \right) \left| \frac{d}{dv} \log \left(\Pi_1^{(j)}(F_N)(v) \right) \right|^2 \Pi_1^{(j)}(F_N)(v) dv \right)^{\frac{q}{2(q-1)}} \right. \\ &\quad \left. \left(P_q^{(j)}(F_N) \right)^{\frac{q-2}{2(q-1)}} \right. \\ &\quad \left. + 1 + M_2 \left(\Pi_1^{(j)}(F_N)(v) \right) \right)^{\frac{q-1}{q}} W_q \left(\Pi_1^{(j)}(F_N), \Pi_1^{(j)}(\widetilde{F}_N) \right). \end{aligned}$$

The result follows from the inequality

$$\int_{\mathbb{R}} \left(1 - \frac{v^2}{N} \right) \left| \frac{d}{dv} \log \left(\Pi_1^{(j)}(F_N)(v) \right) \right|^2 \Pi_1^{(j)}(F_N)(v) dv \leq I_N \left(F_j^{(N)} \right) + 2 \frac{N-3}{N},$$

which is a consequence of Lemma 3.6, and Corollary 3.4. □

4 The Entropy Relation: From Marginals on the Sphere to Marginals on the Line

In Sect. 3 we have seen how to relate the relative entropy of $\Pi_1^{(j)}(F_N)$ to that of $\Pi_1^{(j)}(\widetilde{F}_N)$, gaining a quantitative estimation of the difference between the two. However, our entropic inequalities, (1.9) and (1.10), relate to the entropy of the marginal on the sphere. In this section we will explore the connection between the entropies of the marginals on the sphere and those of the marginals.

Lemma 4.1 *Let $F_N \in P(\mathbb{S}^{N-1}(\sqrt{N}))$. Then*

$$\begin{aligned} \int_{\mathbb{S}^{N-1}(\sqrt{N})} F_j^{(N)} \log F_j^{(N)} d\sigma^N &= H(\Pi_1^{(j)}(F_N) | \gamma) - \log\left(1 - \frac{3}{4N} + o\left(\frac{1}{N}\right)\right) \\ &\quad - \frac{1}{2} \int_{-\sqrt{N}}^{\sqrt{N}} v^2 \Pi_1^{(j)}(F_N)(v) dv - \frac{N-3}{2} \int_{-\sqrt{N}}^{\sqrt{N}} \Pi_1^{(j)}(F_N)(v) \log\left(1 - \frac{v^2}{N}\right) dv. \end{aligned} \tag{4.1}$$

Proof Using (2.3) we find that

$$\begin{aligned} &\int_{\mathbb{S}^{N-1}(\sqrt{N})} F_j^{(N)} \log F_j^{(N)} d\sigma^N \\ &= \frac{|\mathbb{S}^{N-2}|}{|\mathbb{S}^{N-1}|\sqrt{N}} \int_{-\sqrt{N}}^{\sqrt{N}} \left(1 - \frac{v^2}{N}\right)_+^{\frac{N-3}{2}} F_j^{(N)}(v) \log(F_j^{(N)}(v)) dv \\ &= \int_{-\sqrt{N}}^{\sqrt{N}} \Pi_1^{(j)}(F_N)(v) \log(\Pi_1^{(j)}(F_N)(v)) dv \\ &\quad - \int_{-\sqrt{N}}^{\sqrt{N}} \Pi_1^{(j)}(F_N)(v) \log\left(\frac{|\mathbb{S}^{N-2}|}{|\mathbb{S}^{N-1}|\sqrt{N}} \left(1 - \frac{v^2}{N}\right)^{\frac{N-3}{2}}\right) dv \\ &= H(\Pi_1^{(j)}(F_N) | \gamma) + \log\left(\frac{1}{\sqrt{2\pi}}\right) \\ &\quad - \frac{1}{2} \int_{-\sqrt{N}}^{\sqrt{N}} v^2 \Pi_1^{(j)}(F_N)(v) dv - \log\left(\frac{|\mathbb{S}^{N-2}|}{|\mathbb{S}^{N-1}|\sqrt{N}}\right) \\ &\quad - \frac{N-3}{2} \int_{-\sqrt{N}}^{\sqrt{N}} \Pi_1^{(j)}(F_N)(v) \log\left(1 - \frac{v^2}{N}\right) dv, \end{aligned}$$

yielding the desired result. □

Lemma 4.2 *Let $F_N \in P(\mathbb{S}^{N-1}(\sqrt{N}))$ such that $M_k(\Pi_1^{(j)}(F_N)) < \infty$ for some $k > 2$.*

(i) If $I\left(\Pi_1^{(j)}(F_N)\right) < \infty$ then for any sequence $0 < \epsilon_N < 1$, converging to zero, we have that

$$\begin{aligned}
 &-\frac{1}{2} \int_{-\sqrt{N}}^{\sqrt{N}} v^2 \Pi_1^{(j)}(F_N)(v) dv - \frac{N-3}{2} \int_{-\sqrt{N}}^{\sqrt{N}} \Pi_1^{(j)}(F_N)(v) \log\left(1 - \frac{v^2}{N}\right) dv \\
 &\leq \frac{M_k\left(\Pi_1^{(j)}(F_N)\right)}{2N^{\frac{k}{2}-1}\epsilon_N} + \frac{I\left(\Pi_1^{(j)}(F_N)\right)^{\frac{p-1}{2p}} M_k\left(\Pi_1^{(j)}(F_N)\right)^{\frac{1}{p}} C_p}{2(1-\epsilon_N)^{\frac{k}{2p}} N^{\frac{1}{2}\left(\frac{k+1}{p}-3\right)}}
 \end{aligned} \tag{4.2}$$

where $1 < p < \frac{k}{2}$ and

$$C_p = \left(\int_{|x|<1} \left| \log(1-x^2) \right|^{\frac{p}{p-1}} dx \right)^{\frac{p-1}{p}}. \tag{4.3}$$

(ii) If $I_N\left(F_j^{(N)}\right) < \infty$ then for any sequence $0 < \epsilon_N < 1$, converging to zero, we have that

$$\begin{aligned}
 &-\frac{1}{2} \int_{-\sqrt{N}}^{\sqrt{N}} v^2 \Pi_1^{(j)}(F_N)(v) dv - \frac{N-3}{2} \int_{-\sqrt{N}}^{\sqrt{N}} \Pi_1^{(j)}(F_N)(v) \log\left(1 - \frac{v^2}{N}\right) dv \\
 &\leq \frac{M_k\left(\Pi_1^{(j)}(F_N)\right)}{2N^{\frac{k}{2}-1}\epsilon_N} + \frac{N}{2(N-3)(1-\epsilon_N)^{\frac{k}{4}+\frac{1}{2}}} \\
 &\quad \left(I_N\left(F_j^{(N)}\right) + 2\frac{N-3}{N} \right)^{\frac{1}{2}} \frac{l_N}{N^{\frac{k}{4}-\frac{1}{2}}} M_k\left(\Pi_1^{(j)}(F_N)\right)^{\frac{1}{2}},
 \end{aligned} \tag{4.4}$$

where $l_N = \sqrt{\sup_{x \in [0, \epsilon_N]} x (\log x)^2}$

Proof Using the inequality

$$-\log(1-x) < \frac{x}{1-x}$$

for $0 < x < 1$, we find that

$$-\frac{N-3}{2} \log\left(1 - \frac{v^2}{N}\right) - \frac{v^2}{2} < \frac{N-3}{2} \frac{v^2}{N-v^2} - \frac{v^2}{2} < \frac{v^4}{2(N-v^2)}.$$

For any $R > 0$ we have that

$$\begin{aligned}
 &\int_{|v|<R} \left(-\frac{N-3}{2} \log\left(1 - \frac{v^2}{N}\right) - \frac{v^2}{2} \right) \Pi_1^{(j)}(F_N)(v) dv \\
 &\leq \frac{1}{2(N-R^2)} \int_{|v|<R} v^4 \Pi_1^{(j)}(F_N)(v) dv \leq \frac{R^{4-k}}{2(N-R^2)} M_k\left(\Pi_1^{(j)}(F_N)\right).
 \end{aligned}$$

Picking $R = \sqrt{N(1 - \epsilon_N)}$, with $0 < \epsilon_N < 1$ going to zero, we find that

$$\int_{|v| < \sqrt{N(1 - \epsilon_N)}} \left(-\frac{N - 3}{2} \log \left(1 - \frac{v^2}{N} \right) - \frac{v^2}{2} \right) \Pi_1^{(j)}(F_N)(v) dv \leq \frac{M_k \left(\Pi_1^{(j)}(F_N) \right)}{2N^{\frac{k}{2} - 1} \epsilon_N}. \tag{4.5}$$

The difference between (i) and (ii) manifests itself in the domain $|v| \geq \sqrt{N(1 - \epsilon_N)}$. To prove (i) we notice that

$$\begin{aligned} & -\frac{1}{2} \int_{|v| \geq \sqrt{N(1 - \epsilon_N)}} v^2 \Pi_1^{(j)}(F_N)(v) dv \\ & - \frac{N - 3}{2} \int_{|v| \geq \sqrt{N(1 - \epsilon_N)}} \Pi_1^{(j)}(F_N)(v) \log \left(1 - \frac{v^2}{N} \right) dv \\ & \leq -\frac{1}{2} \int_{|v| \geq \sqrt{N(1 - \epsilon_N)}} N \Pi_1^{(j)}(F_N)(v) \log \left(1 - \frac{v^2}{N} \right) dv \\ & \leq \frac{1}{2(1 - \epsilon_N)} \int_{|v| \geq \sqrt{N(1 - \epsilon_N)}} v^2 \Pi_1^{(j)}(F_N)(v) \left(-\log \left(1 - \frac{v^2}{N} \right) \right) dv \\ & \leq \frac{1}{2(1 - \epsilon_N)} \left(\int_{|v| \geq \sqrt{N(1 - \epsilon_N)}} |v|^{2p} \Pi_1^{(j)}(F_N)(v) dv \right)^{\frac{1}{p}} \\ & \left(\int_{-\sqrt{N}}^{\sqrt{N}} \left| \log \left(1 - \frac{v^2}{N} \right) \right|^{\frac{p-1}{p}} \Pi_1^{(j)}(F_N)(v) dv \right)^{\frac{p-1}{p}} \\ & \leq \frac{1}{2(1 - \epsilon_N)} \left(\frac{1}{(N(1 - \epsilon_N))^{\frac{k}{2} - p}} \int_{|v| \geq \sqrt{N(1 - \epsilon_N)}} |v|^k \Pi_1^{(j)}(F_N)(v) dv \right)^{\frac{1}{p}} \\ & \left\| \Pi_1^{(j)}(F_N) \right\|_{\infty}^{\frac{p-1}{p}} \left(\int_{-\sqrt{N}}^{\sqrt{N}} \left| \log \left(1 - \frac{v^2}{N} \right) \right|^{\frac{p-1}{p}} dv \right)^{\frac{p-1}{p}} \\ & = \frac{\left\| \Pi_1^{(j)}(F_N) \right\|_{\infty}^{\frac{p-1}{p}} M_k \left(\Pi_1^{(j)}(F_N) \right)^{\frac{1}{p}} C_p}{2(1 - \epsilon_N)^{\frac{k}{2p}} N^{\frac{1}{2} \left(\frac{k+1}{p} - 3 \right)}}, \end{aligned} \tag{4.6}$$

where $p > 1$ was chosen such that $p < \frac{k}{2}$. The result follows from (4.5), (4.6) and the following inequality: For any $f \in P(\mathbb{R})$ with a finite Fisher information $I(f)$ one has that

$$\|f\|_{\infty} \leq (I(f))^{\frac{1}{2}}.$$

In order to prove (ii) we notice that

$$\begin{aligned}
 & -\frac{1}{2} \int_{|v| \geq \sqrt{N(1-\epsilon_N)}} v^2 \Pi_1^{(j)}(F_N)(v) dv \\
 & \quad - \frac{N-3}{2} \int_{|v| \geq \sqrt{N(1-\epsilon_N)}} \Pi_1^{(j)}(F_N)(v) \log\left(1 - \frac{v^2}{N}\right) dv \\
 & \leq \frac{1}{2(1-\epsilon_N)} \int_{|v| \geq \sqrt{N(1-\epsilon_N)}} v^2 \Pi_1^{(j)}(F_N)(v) \left| \log\left(1 - \frac{v^2}{N}\right) \right| dv \\
 & \leq \frac{1}{2(1-\epsilon_N)} \left(\int_{|v| \geq \sqrt{N(1-\epsilon_N)}} \frac{v^2 \Pi_1^{(j)}(F_N)(v)}{1 - \frac{v^2}{N}} \right)^{\frac{1}{2}} \\
 & \quad \left(\int_{|v| \geq \sqrt{N(1-\epsilon_N)}} v^2 \left(1 - \frac{v^2}{N}\right) \left| \log\left(1 - \frac{v^2}{N}\right) \right|^2 \Pi_1^{(j)}(F_N)(v) dv \right)^{\frac{1}{2}} \\
 & \leq \frac{N}{2(N-3)(1-\epsilon_N)} \left(I_N(F_j^{(N)}) + 2 \frac{N-3}{N} \right)^{\frac{1}{2}} \\
 & \quad \frac{I_N}{(N(1-\epsilon_N))^{\frac{k}{4}-\frac{1}{2}}} M_k \left(\Pi_1^{(j)}(F_N) \right)^{\frac{1}{2}},
 \end{aligned}$$

showing the result. □

Combining Lemmas 4.1 and 4.2 with the choice $\epsilon_N = N^{-\beta}$ gives us

Theorem 4.3 *Let $F_N \in P\left(\mathbb{S}^{N-1}\left(\sqrt{N}\right)\right)$ such that $M_k\left(\Pi_1^{(j)}\left(F_N\right)\right) < \infty$ for some $k > 2$.*

(i) If $I\left(\Pi_1^{(j)}\left(F_N\right)\right) < \infty$ then there exists $C_1 > 0$, independent of N and F_N , such that for any $\beta > 0$ and any $1 < p < \min\left(\frac{k+1}{3}, \frac{k}{2}\right)$

$$\begin{aligned}
 & \int_{\mathbb{S}^{N-1}\left(\sqrt{N}\right)} F_j^{(N)} \log F_j^{(N)} d\sigma^N \leq H\left(\Pi_1^{(j)}\left(F_N\right) | \gamma\right) + \frac{C_1}{N} \\
 & \quad + \frac{M_k\left(\Pi_1^{(j)}\left(F_N\right)\right)}{2N^{\frac{k}{2}-1-\beta}} + \frac{I\left(\Pi_1^{(j)}\left(F_N\right)\right)^{\frac{p-1}{2p}} M_k\left(\Pi_1^{(j)}\left(F_N\right)\right)^{\frac{1}{p}} C_p}{2\left(1 - \frac{1}{N^\beta}\right)^{\frac{k}{2p}} N^{\frac{1}{2}\left(\frac{k+1}{p}-3\right)}} \tag{4.7} \\
 & = H\left(\Pi_1^{(j)}\left(F_N\right) | \gamma\right) + \tilde{\epsilon}_N^{1,j},
 \end{aligned}$$

where $C_p = \left(\int_{|x|<1} |\log(1-x^2)|^{\frac{p}{p-1}} dx\right)^{\frac{p-1}{p}}$.

(ii) If $I_N \left(F_j^{(N)} \right) < \infty$ then there exists $C_1 > 0$, independent of N and F_N , such that for any $\beta > 0$

$$\begin{aligned} \int_{\mathbb{S}^{N-1}(\sqrt{N})} F_j^{(N)} \log F_j^{(N)} d\sigma^N &\leq H \left(\Pi_1^{(j)} (F_N) \mid \gamma \right) + \frac{C_1}{N} \\ &\leq \frac{M_k \left(\Pi_1^{(j)} (F_N) \right)}{2N^{\frac{k}{2}-1-\beta}} + \frac{N}{2(N-3)\left(1 - \frac{1}{N^\beta}\right)^{\frac{k}{4}+\frac{1}{2}}} \\ &\quad \left(I_N \left(F_j^{(N)} \right) + 2\frac{N-3}{N} \right)^{\frac{1}{2}} \frac{\eta_{N,\beta}}{N^{\frac{k}{4}-\frac{1}{2}}} M_k \left(\Pi_1^{(j)} (F_N) \right)^{\frac{1}{2}} \\ &= H \left(\Pi_1^{(j)} (F_N) \mid \gamma \right) + \tilde{\epsilon}_N^{1,j}, \end{aligned} \tag{4.8}$$

where $\eta_{N,\beta} = \sqrt{\sup_{x \in [0, N-\beta]} x (\log x)^2}$.

We now have all the tools to prove our main theorems.

5 Proof of the Main Theorems

In the previous couple of sections we have managed to find conditions on our original probability density, F_N , such that the appropriate marginals on the sphere, marginals on the line and the marginals of the extension give close values for the appropriate entropy functional. Combining all these results will lead to the proof of our main theorems, which is the subject of this section.

We begin with a simple technical lemma, whose proof is a simple application of the Hölder inequality:

Lemma 5.1 Let $\{a_{j,i}\}_{j=1,\dots,m \ i=1,\dots,N}$ be non-negative numbers. Let p_1, \dots, p_m be positive numbers such that $\sum_{j=1}^m \frac{1}{p_j} \leq 1$. Then

$$\sum_{i=1}^N \left(\prod_{j=1}^m a_{j,i}^{\frac{1}{p_j}} \right) \leq \prod_{j=1}^m \left(\frac{\sum_{i=1}^N a_{j,i}}{N} \right)^{\frac{1}{p_j}} N. \tag{5.1}$$

Theorem 5.2 Let $F_N \in P \left(\mathbb{S}^{N-1} \left(\sqrt{N} \right) \right)$ such that there exists $k > 2$ with

$$\mathcal{A}_{N,k}^M = \frac{\sum_{j=1}^N M_k \left(\Pi_1^{(j)} (F_N) \right)}{N} < \infty.$$

Assume in addition that

$$\mathcal{A}_N^I = \frac{\sum_{i=1}^N I \left(\Pi_1^{(j)} (F_N) \right)}{N} < \infty.$$

Then there exists $C_1, C_2 > 0$ independent of N and F_N , such that for any $\beta > 0$ and $1 < p < \min\left(\frac{k+1}{3}, \frac{k}{2}\right)$

$$\begin{aligned} & \sum_{j=1}^N \int_{\mathbb{S}^{N-1}(\sqrt{N})} F_j^{(N)} \log F_j^{(N)} d\sigma^N \leq H_N(F_N) + C_1 \\ & + \left(\frac{4C_2 \left(1 + \frac{C_k}{2}\right)^{\frac{1}{k}}}{(2N)^{\frac{1}{4} - \frac{1}{2k}}} \left(\mathcal{A}_N^I - 1\right)^{\frac{1}{2}} \left(1 + \mathcal{A}_{N,k}^M\right)^{\frac{1}{k}} \right) N \\ & + \left(\frac{\mathcal{A}_{N,k}^M}{2N^{\frac{k}{2} - 1 - \beta}} \right) N + \left(\frac{C_p \left(\mathcal{A}_N^I\right)^{\frac{p-1}{2p}} \left(\mathcal{A}_{N,k}^M\right)^{\frac{1}{p}}}{2 \left(1 - \frac{1}{N^\beta}\right)^{\frac{k}{2p}} N^{\frac{1}{2} \left(\frac{k+1}{p} - 3\right)}} \right) N, \end{aligned} \tag{5.2}$$

where $C_p = \left(\int_{|x|<1} |\log(1-x^2)|^{\frac{p}{p-1}}\right)^{\frac{p-1}{p}}$ and $C_k = \sup_N \left(\frac{2}{N}\right)^{\frac{k}{2}} \frac{\Gamma\left(\frac{N+k}{2}\right)}{\Gamma\left(\frac{N}{2}\right)}$.

Proof This follows immediately from Theorems 3.7, 4.3, and Lemma 5.1, the fact that for any $F_N \in P\left(\mathbb{S}^{N-1}(\sqrt{N})\right)$

$$\sum_{j=1}^N M_2\left(\Pi_1^{(j)}(F_N)\right) = \sum_{j=1}^N \int_{\mathbb{S}^{N-1}(\sqrt{N})} v_j^2 F_N d\sigma^N = N,$$

and inequality (1.2) applied to \widetilde{F}_N together with

$$H\left(\widetilde{F}_N | \gamma_N\right) = H_N(F_N),$$

proven in Lemma 2.2. □

Theorem 5.3 Let $F_N \in P\left(\mathbb{S}^{N-1}(\sqrt{N})\right)$ such that there exists $k > 2$ with

$$\mathcal{A}_{N,k}^M = \frac{\sum_{j=1}^N M_k\left(\Pi_1^{(j)}(F_N)\right)}{N} < \infty.$$

Assume in addition that

$$\mathcal{A}_N^{I_S} = \frac{\sum_{i=1}^N I_N\left(\Pi_1^{(j)}(F_N)\right)}{N} < \infty,$$

and that there exists $2 < q < k$ such that

$$\mathcal{A}_{N,q}^P = \frac{\sum_{j=1}^N P_q^{(j)}(F_N)}{N} < \infty,$$

where

$$P_q^{(j)}(F_N) = \int \frac{\Pi_1^{(j)}(F_N)(v)}{\left(1 - \frac{v^2}{N}\right)^{\frac{q}{q-2}}} dv.$$

Then there exists $C_1, C_2 > 0$ independent of N and F_N , such that for any $\beta > 0$

$$\begin{aligned} \sum_{j=1}^N \int_{\mathbb{S}^{N-1}(\sqrt{N})} F_j^{(N)} \log F_j^{(N)} d\sigma^N &\leq H_N(F_N) + C_1 \\ &+ C_2 2^{\frac{3}{2} + \frac{2}{q}} \left(1 + \frac{C_k}{2}\right)^{\frac{1}{k}} \left(\left(\mathcal{A}_N^{I_S} + 2\right)^{\frac{q}{2(q-1)}} \left(\mathcal{A}_{N,q}^P\right)^{\frac{q-2}{2(q-1)}} + 2 \right)^{\frac{q}{q-1}} \frac{\left(1 + \mathcal{A}_{N,k}^M\right)^{\frac{1}{k}}}{(2N)^{\frac{1}{2q} - \frac{1}{2k}}} N \\ &+ \left(\frac{\mathcal{A}_{N,k}^M}{2N^{\frac{k}{2} - 1 - \beta}}\right) N + \frac{N}{2(N-3)} \frac{\eta_{N,\beta}}{N^{\frac{k}{4} - \frac{1}{2}} \left(1 - \frac{1}{N^\beta}\right)^{\frac{k}{4} + \frac{1}{2}}} \left(\mathcal{A}_N^{I_S} + 2\right)^{\frac{1}{2}} \left(\mathcal{A}_{N,k}^M\right)^{\frac{1}{2}} N, \end{aligned} \tag{5.3}$$

where $C_k = \sup_N \left(\frac{2}{N}\right)^{\frac{k}{2}} \frac{\Gamma\left(\frac{N+k}{2}\right)}{\Gamma\left(\frac{N}{2}\right)}$ and $\eta_{N\beta} = \sup_{x \in [0, N-\beta]} x (\log x)^2$.

Proof Much like the proof of Theorem 5.2, we just rely on Theorems 3.7, 4.3, Lemma 5.1, the simple second moment computation and the entropic inequality for \widetilde{F}_N . \square

Proof of Theorem 1.7 This follows immediately from Theorem 5.2 and the fact that

$$N \leq \frac{H_N(F_N)}{C_H}.$$

\square

Proof of Theorem 1.8 This follows immediately from Theorem 5.3, the known inequality

$$\sum_{j=1}^N I_N \left(F_j^{(N)}\right) \leq 2I_N(F_N) \tag{5.4}$$

(see [1]) and, much like the proof of Theorem 1.7, the fact that $N \leq \frac{H_N(F_N)}{C_H}$. \square

6 A Non-trivial Example

As was mentioned in the Introduction of this work, there is a connection between inequality (1.4) and the subject of entropic convergence to equilibrium in Kac’s model (the many-body Cercignani’s conjecture). It is thus not surprising that in order to find a family of density functions that will serve as an example to the validity of the conditions of our main theorems, we look for natural “states” occurring in the setting of Kac’s model. Such states, intimately connected to the concept of *chaoticity* and

entropic chaoticity, are described below (for more information, we refer the reader to [4,5,8,10,11]).

Given $f \in P(\mathbb{R})$, with additional conditions we will mention shortly, we can define the normalization function, $\mathcal{Z}_N(f, r)$, as

$$\mathcal{Z}_N(f, r) = \int_{\mathbb{S}^{N-1}(r)} f^{\otimes N} d\sigma_r^N.$$

The conditioned tensorization of f on the sphere is the probability measure on $\mathbb{S}^{N-1}(\sqrt{N})$ with density

$$F_N = \frac{f^{\otimes N}}{\mathcal{Z}_N(f, \sqrt{N})}.$$

The following theorem, proved in [4], is of great importance in the study of conditioned tensorizations, and reinforces the intuition that when f has a unit second moment the N -tensorization function of f is concentrated tightly about $\mathbb{S}^{N-1}(\sqrt{N})$.

Theorem 6.1 *Let $f \in P(\mathbb{R})$ such that $f \in L^p(\mathbb{R})$ for some $p > 1$, $\int_{\mathbb{R}} v^2 f(v)dv = 1$ and $\int_{\mathbb{R}} v^4 f(v)dv < \infty$. Then*

$$\mathcal{Z}_N(f, \sqrt{u}) = \frac{2}{\sqrt{N}\Sigma |\mathbb{S}^{N-1}| u^{\frac{N-2}{2}}} \left(\frac{e^{-\frac{(u-N)^2}{2N\Sigma^2}}}{\sqrt{2\pi}} + \lambda_N(u) \right), \tag{6.1}$$

where $\Sigma^2 = \int_{\mathbb{R}} v^4 f(v)dv - 1$ and $\sup_u |\lambda_N(u)| \xrightarrow{N \rightarrow \infty} 0$.

We are now ready to present our non-trivial example for a family of densities on the sphere that satisfies the conditions of our main theorems. While extensions of it can be found, we restrict ourselves to a relatively simple case to avoid some lengthy computations.

Theorem 6.2 *Let $f \in P(\mathbb{R}) \cap C_c(\mathbb{R})$, $f \neq \gamma$, be such that $\int_{\mathbb{R}} v^2 f(v) = 1$ and $I(f) < \infty$. Then, the conditioned tensorization of f satisfies the conditions of Theorems 1.7 and 1.8.*

Proof The first thing we note is that since F_N is symmetric with respect to its variables all the marginals are identical. As such, for any $j \geq 2$

$$H_N \left(\Pi_1^{(j)}(F_N) \right) = H_N \left(\Pi_1^{(1)}(F_N) \right) = H_N \left(\Pi_1(F_N) \right)$$

and the same holds for I, I_N and M_k . In that case the appropriate averaged quantities, \mathcal{A} , are

$$\mathcal{A}_k = \sup_N M_k \left(\Pi_1(F_N) \right), \quad \mathcal{A}_I = \sup_N I \left(\Pi_1(F_N) \right), \quad \mathcal{A}_q^P = \sup_N P_q^{(1)} \left(\Pi_1(F_N) \right).$$

Using formula (2.4) and the definition of the normalization function we have that

$$\begin{aligned} \Pi_1(F_N)(v) &= \frac{|\mathbb{S}^{N-2}| \left(1 - \frac{v^2}{N}\right)^{\frac{N-3}{2}} \mathcal{Z}_{N-1}\left(f, \sqrt{N-v^2}\right)}{|\mathbb{S}^{N-1}| \sqrt{N} \mathcal{Z}_N\left(f, \sqrt{N}\right)} f(v) \\ &= \sqrt{\frac{N}{N-1}} \frac{e^{-\frac{(1-v^2)^2}{2(N-1)\Sigma^2}} + \sqrt{2\pi}\lambda_{N-1}(N-v^2)}{1 + \sqrt{2\pi}\lambda_N(N)} f(v), \end{aligned} \tag{6.2}$$

due to (6.1). As such

$$\mathcal{A}_k = \sup_N M_k(\Pi_1(F_N)) \leq \sup_N \frac{1 + \sqrt{2\pi} \sup |\lambda_{N-1}|}{1 + \sqrt{2\pi}\lambda_N(N)} \sqrt{\frac{N}{N-1}} \int_{\mathbb{R}} |v|^k f(v) dv < \infty,$$

for any $k > 0$ as $f \in C_c(\mathbb{R})$.

Let $R > 0$ be such that f is supported in $[-R, R]$. We find that for $N > R$

$$\begin{aligned} \mathcal{A}_q^P &= \sup_N P_q^{(1)}(\Pi_1(F_N)) = \sup_N \int_{-R}^R \frac{\Pi_1(F_N)(v)}{\left(1 - \frac{v^2}{N}\right)^{\frac{q}{q-2}}} dv \\ &\leq \sup_N \frac{1 + \sqrt{2\pi} \sup |\lambda_{N-1}|}{1 + \sqrt{2\pi}\lambda_N(N)} \sqrt{\frac{N}{N-1}} \frac{1}{\left(1 - \frac{R^2}{N}\right)^{\frac{q}{q-2}}} < \infty, \end{aligned}$$

for any $q > 2$.

Using (3.7) and the fact that f is compactly supported, we see that for $N > R$

$$\begin{aligned} I(\Pi_1(F_N)) &= \int_{-R}^R \left| \frac{d}{dv} \log \Pi_1(F_N)(v) \right| \Pi_1(F_N)(v) dv \\ &\leq \frac{I_N(F_1^{(N)}) + 2}{\left(1 - \frac{R^2}{N}\right)} \leq \frac{2\left(\frac{I_N(F_N)}{N} + 1\right)}{\left(1 - \frac{R^2}{N}\right)}, \end{aligned}$$

where we have used (5.4) and the symmetry of F_N . This implies that

$$\mathcal{A}_I = \sup_N I(\Pi_1(F_N)) \leq \sup_N \frac{2\left(\frac{I_N(F_N)}{N} + 1\right)}{\left(1 - \frac{R^2}{N}\right)},$$

showing that if

$$\sup_N \frac{I_N(F_N)}{N} < \infty$$

we obtain the required Fisher information condition for Theorem 1.7, as well as Theorem 1.8. We find that

$$\begin{aligned} \frac{I_N(F_N)}{N} &= \frac{1}{N} \int_{\mathbb{S}^{N-1}(\sqrt{N})} \frac{|\nabla_{\mathbb{S}} F_N|^2}{F_N} d\sigma^N \leq \frac{1}{N} \int_{\mathbb{S}^{N-1}(\sqrt{N})} \frac{|\nabla F_N|^2}{F_N} d\sigma^N \\ &= \int_{\mathbb{S}^{N-1}(\sqrt{N})} \left(\frac{f'(v_1)}{f(v_1)} \right)^2 F_N d\sigma^N = \int_{\mathbb{R}} \left(\frac{f'(v)}{f(v)} \right)^2 \Pi_1(F_N)(v) dv \\ &\leq \sup_N \frac{1 + \sqrt{2\pi} \sup |\lambda_{N-1}|}{1 + \sqrt{2\pi} \lambda_N(N)} \sqrt{\frac{N}{N-1}} I(f) = C_I < \infty, \end{aligned}$$

where we have used the special structure of F_N and symmetry.

Last, but not least, we will deal with the rescaled entropy term.

$$\begin{aligned} \frac{H_N(F_N)}{N} &= \frac{1}{N} \int_{\mathbb{S}^{N-1}(\sqrt{N})} F_N \log f^{\otimes N} d\sigma^N - \frac{\log \mathcal{Z}_N(f, \sqrt{N})}{N} \\ &= \int_{\mathbb{R}} \log f(v) \Pi_1(F_N)(v) dv + \frac{\log \left(|\mathbb{S}^{N-1}| N^{\frac{N-1}{2}} \right)}{N} \\ &\quad - \frac{\log \left(\frac{2}{\sqrt{2\pi} \Sigma} \left(1 + \sqrt{2\pi} \lambda_N(N) \right) \right)}{N}. \end{aligned}$$

As f is supported on $[-R, R]$ we find that $\Pi_1(F_N)$ converges to f uniformly on \mathbb{R} . Also, using the asymptotic approximation of $|\mathbb{S}^{N-1}|$ one can show that

$$\frac{\log \left(|\mathbb{S}^{N-1}| N^{\frac{N-1}{2}} \right)}{N} \xrightarrow{N \rightarrow \infty} \frac{1 + \log(2\pi)}{2} = - \int_{\mathbb{R}} f(v) \log \gamma(v) dv.$$

Thus,

$$\lim_{N \rightarrow \infty} \frac{H_N(F_N)}{N} = H(f|\gamma) > 0,$$

and since $F_N \not\equiv 1$ we know that $H_N(F_N) \neq 0$ for all N , implying that there exists $C_H > 0$ with

$$\frac{H_N(F_N)}{N} \geq C_H,$$

completing our theorem. □

Remark 6.3 Note that in the proof of the above theorem the only quantity that wasn't bounded by an "explicit" constant is the rescaled entropy. However, such a constant can be found by a more detailed computation.

7 Final Remarks

While the main result proved in this paper gives a glimpse of tools and quantities that are of import both to the equivalence of ensembles and the many-body Cercignani's conjecture, there are still many items of interest that can be explored in future research. We present a few remarks and observations related to that:

- The condition on the pole control, $P_q^{(j)}$, seems to fit the problematic behavior near the poles that was used to show that the constant in (1.4) is sharp. However, in relation to Kac's model, it seems hard to show the propagation of such a property under Kac's flow. If one is allowed to use the exponent $q = \infty$, it is easy to see that the expression given for $P_\infty^{(j)}$ is controlled by $I_N(F_j^{(N)})$ —a more natural quantity in the kinetic setting. It would be interesting to see what will need to replace, if possible, the condition about infinite moment control (i.e., $k = \infty$) in order to be able to use this.
- The moment control condition appears to be natural in Kac's setting. Indeed, following [7] one sees that the family of functions that was constructed to show the validity of Villani's conjecture satisfies

$$M_k(\Pi_1(F_N)) \xrightarrow{N \rightarrow \infty} \infty,$$

for any $k > 2$.

- A very important observation that can be made following Theorems 5.2 and 5.3 is that the requirement on $\frac{H_N(F_N)}{N}$ can be removed and one can gain a quantitative version of the deviation of the sum of the partial entropies with respect to the total entropy. In other words, we can find an explicit κ_N such that

$$\sum_{j=1}^N \int_{\mathbb{S}^{N-1}(\sqrt{N})} F_j^{(N)} \log F_j^{(N)} d\sigma^N \leq H_N(F_N) + \kappa_N.$$

Under our setting κ_N may blow up, but perhaps a more delicate estimation can be done in the future to evaluate it, or some regimes on the behavior of $H_N(F_N)$ may be explored and will allow us to improve our main inequality.

- The rescaled entropy, $\frac{H_N(F_N)}{N}$ is very important in the study of Kac's model and is connected to the concept of entropic chaoticity (see more in [4, 9–11]). One knows that under Kac's flow the entropy will decrease, so a lower bound on the rescaled entropy can't propagate with time. However, it may give rise to a two time scale approach where we find a fast convergence to a state near equilibrium if we start far from equilibrium using the ideas in our work, followed by a fast convergence to equilibrium using different techniques.

Acknowledgments We would like to greatly thank Eric Carlen for many discussions and insights on key ideas all along the progression of this work, without which this paper would never have seen the light of day. We would also like to offer our gratitude to Nathael Gozlan for providing us with a reference for the "distorted" HWI inequality we use in Sect. 3, and Clément Mouhot for several discussions on the present results. The author was supported by EPSRC Grant EP/L002302/1.

Appendix: Additional Proofs

In this Appendix we will provide additional proofs that we felt would hinder the flow of the paper.

Lemma 8.1 *Let $F_N \in P(\mathbb{R}^N)$ be a probability density with finite second moment. Then*

$$\sum_{j=1}^N H\left(\Pi_1^{(j)}(F_N) \mid \gamma\right) \leq H(F_N \mid \gamma_N). \tag{8.1}$$

Proof It is a simple computation to see that

$$\sum_{j=1}^N \int_{\mathbb{R}^N} \Pi_1^{(j)}(F_N)(v_j) \log \gamma(v_j) dv_j = \int_{\mathbb{R}^N} F_N(v) \log \gamma_N(v) dv.$$

Thus, we only need to prove that

$$\sum_{j=1}^N H\left(\Pi_1^{(j)}(F_N)\right) \leq H(F_N).$$

Define $G_N(v) = \prod_{j=1}^N \Pi_1^{(j)}(F_N)(v_j)$. $G_N \in P(\mathbb{R}^N)$ and

$$\begin{aligned} 0 \leq H(F_N \mid G_N) &= H(F_N) - \int_{\mathbb{R}^N} F_N(v) \log G_N(v) dv \\ &= H(F_N) - \sum_{j=1}^N \int_{\mathbb{R}^N} F_N(v) \log \left(\Pi_1^{(j)}(F_N)(v_j)\right) dv \\ &= H(F_N) - \sum_{j=1}^N H\left(\Pi_1^{(j)}(F_N)\right), \end{aligned}$$

completing the proof. □

Lemma 8.2 *Denote by $d(x, y) = \min(|x - y|, 1)$ for any $x, y \in \mathbb{R}$. Then for any $q \geq 1$ and $R \geq 1$*

$$|x - y|^q \leq R^q d(x, y) + \frac{2^k}{R^{k-q}} \left(|x|^k + |y|^k\right). \tag{8.2}$$

Proof If $|x - y| \leq 1$ we have that

$$|x - y|^q \leq |x - y| = d(x, y) \leq R^q d(x, y) + \frac{2^k}{R^{k-q}} \left(|x|^k + |y|^k\right).$$

When $|x - y| > 1$ we have that if $|x|, |y| < \frac{R}{2}$

$$\begin{aligned} |x - y|^q &\leq 2^{q-1} (|x|^q + |y|^q) \leq R^q = R^q d(x, y) \\ &\leq R^q d(x, y) + \frac{2^k}{R^{k-q}} (|x|^k + |y|^k), \end{aligned}$$

due to the convexity of the map $f(t) = t^q$. If $|x| < \frac{R}{2}$ and $|y| > \frac{R}{2}$ (or vice versa)

$$\begin{aligned} |x - y|^q &\leq 2^{q-1} (|x|^q + |y|^q) \leq \frac{R^q}{2} + 2^{q-1} \left(\frac{2}{R}\right)^{k-q} |y|^k \\ &= \frac{R^q}{2} d(x, y) + \frac{2^{k-1}}{R^{k-q}} |y|^k \leq R^q d(x, y) + \frac{2^k}{R^{k-q}} (|x|^k + |y|^k). \end{aligned}$$

Finally, if $|x|, |y| \geq \frac{R}{2}$ then

$$\begin{aligned} |x - y|^q &\leq 2^{q-1} (|x|^q + |y|^q) \leq 2^{q-1} \left(\frac{2}{R}\right)^{k-q} (|y|^k + |x|^k) \\ &\leq R^q d(x, y) + \frac{2^k}{R^{k-q}} (|x|^k + |y|^k), \end{aligned}$$

completing the proof. □

References

1. Barthe, F., Cordero-Erausquin, D., Maurey, B.: Entropy of spherical marginals and related inequalities. *J. Math. Pures Appl.* (9) **86**(2), 89–99 (2006)
2. Carlen, E.A.: The Rate of Local Equilibration in Kinetic Theory, *Prospects in Mathematical Physics*, pp. 71–88, *Contemp. Math.*, 437
3. Carlen, E.A., Lieb, E.H., Loss, M.: A sharp analog of Young’s inequality on S^N and related entropy inequalities. *J. Geom. Anal.* **14**(3), 487–520 (2004)
4. Carlen, E.A., Carvalho, M.C., Le Roux, J., Loss, M., Villani, C.: Entropy and Chaos in the Kac model. *Kinet. Relat. Models* **3**(1), 85–122 (2010)
5. Carrapatoso, K.: Quantitative and Qualitative Kac’s Chaos on the Boltzmann Sphere. [arXiv:1205.1241](https://arxiv.org/abs/1205.1241)
6. Cordero-Erausquin, D.: Some applications of mass transport to Gaussian type inequalities. *Arch. Ration. Mech. Anal.* **161**, 257–269 (2002)
7. Einav, A.: On Villani’s conjecture concerning entropy production for the Kac master equation. *Kinet. Relat. Models* **4**(2), 479–497 (2011)
8. Einav, A.: A counter example to Cercignani’s conjecture for the d -dimensional Kac model. *J. Stat. Phys.* **148**(6), 1076–1103 (2012)
9. Einav, A.: A few ways to destroy entropic chaoticity on Kac’s sphere. *Commun. Math. Sci.* **12**(1), 41–60 (2014)
10. Hauray, M., Mischler, S.: On Kac’s Chaos and related problems. <http://hal.archives-ouvertes.fr/hal-00682782/>
11. Mischler, S., Mouhot, C.: Kac’s program in kinetic theory. *Invent. Math.* **193**(1), 1–147 (2013)
12. Villani, C.: *Topics in Optimal Transportation*, Graduate Studies in Mathematics, vol. 58. American Mathematical Society, Providence (2003)
13. Villani, C.: *Optimal Transport, Old and New*, Grundlehren der Mathematischen Wissenschaften, vol. 338. Springer, London (2009)
14. Villani, C.: Cercignani’s conjecture is sometimes true and always almost true. *Commun. Math. Phys.* **234**(3), 455–490 (2003)