

Chaos for rescaled measures on Kac’s sphere*

Roberto Cortez[†] Hagop Tossounian[‡]

Abstract

In this article we study a relatively novel way of constructing chaotic sequences of probability measures supported on Kac’s sphere, which are obtained as the law of a vector of N i.i.d. variables after it is rescaled to have unit average energy. We show that, as N increases, this sequence is chaotic in the sense of Kac, with respect to the Wasserstein distance, in L^1 , in the entropic sense, and in the Fisher information sense. For many of these results, we provide explicit rates of polynomial order in N . In the process, we improve a quantitative entropic chaos result of Haurey and Mischler by relaxing the finite moment requirement on the densities from order 6 to $4 + \epsilon$.

Keywords: Kac’s chaos; propagation of chaos; entropy; entropic chaos; Fisher information.

MSC2020 subject classifications: 82C40; 60F99.

Submitted to EJP on April 14, 2022, final version accepted on June 2, 2023.

Supersedes arXiv:2204.05406.

1 Introduction

1.1 Chaotic sequences

In this article we study a relatively novel way of constructing chaotic sequences supported on Kac’s sphere. We are largely motivated by the work of Carlen, Carvalho, Le Roux, Loss, and Villani [3]. In this setting, “chaos” is to be understood as “asymptotic statistical independence”, when a parameter $N \in \mathbb{N}$, representing the number of particles, goes to infinity.

More precisely: if $\mathcal{P}(E)$ denotes the space of probability measures on the metric space E , and $\mathcal{P}_{\text{sym}}(\mathbb{R}^N)$ the space of symmetric probability measures on \mathbb{R}^N , that is, those invariant under any permutation of the variables x_1, \dots, x_N , then chaos is defined as follows:

*R. Cortez was supported by ANID Fondecyt Iniciación Grant 11181082. H. Tossounian was supported by ANID Fondecyt Postdoctoral Grant 3200130, and Centro de Modelamiento Matemático (CMM) BASAL fund FB210005 for center of excellence from ANID-Chile.

[†]Universidad Andrés Bello, Departamento de Matemáticas, Santiago, Chile.

E-mail: roberto.cortez.m@unab.cl

[‡]Universidad de Chile, DIM-CMM, Santiago, Chile. E-mail: htossounian@cmm.uchile.cl

Definition 1 (Kac's chaos). For each $N \in \mathbb{N}$, let $F^N \in \mathcal{P}_{\text{sym}}(\mathbb{R}^N)$, and let $f \in \mathcal{P}(\mathbb{R})$. The sequence $(F^N)_{N \in \mathbb{N}}$ is said to be Kac chaotic to f , or simply f -chaotic, if for all $k \in \mathbb{N}$, the projection of F^N on its first k variables (any k variables) converges weakly to $f^{\otimes k}$ as $N \rightarrow \infty$. That is, for any $\phi : \mathbb{R}^k \rightarrow \mathbb{R}$ continuous and bounded, one has

$$\lim_{N \rightarrow \infty} \int_{\mathbb{R}^N} \phi(x_1, \dots, x_k) F^N(dx) = \int_{\mathbb{R}^k} \phi(x_1, \dots, x_k) f(dx_1) \cdots f(dx_k).$$

Kac's sphere for N particles is the $(N-1)$ -dimensional sphere of radius \sqrt{N} embedded in \mathbb{R}^N , that is,

$$\mathcal{K}^N = \left\{ x \in \mathbb{R}^N : \sum_{i=1}^N x_i^2 = N \right\}.$$

It corresponds to the set of average energy equal to 1. Its importance comes from the fact that \mathcal{K}^N is the natural state space of Kac's N -particle system: a pure-jump Markov process on \mathbb{R}^N representing the evolution of the one-dimensional velocities of N identical particles subjected to random energy-preserving collisions; it is a simplification of the particle system associated with the spatially homogeneous Boltzmann equation for dilute gases. In his celebrated paper [11], Kac proved that this model satisfies what is now known as *propagation of chaos*: if the distribution of the system is chaotic at $t = 0$, then it is also chaotic for later times $t > 0$. This provides a bridge between the detailed microscopic description of the gas, given by the particle system, and its reduced macroscopic behaviour, given by the so-called *Boltzmann-Kac equation*. We refer the reader to [3, 13, 15] for more information about Kac's model, the Boltzmann equation, and the important problem of propagation of chaos.

It is then natural to study chaotic sequences of distributions supported on \mathcal{K}^N . An important and archetypal example is the uniform distribution on \mathcal{K}^N , denoted σ^N , which is the unique equilibrium distribution of Kac's N -particle system. As it is well known, σ^N is chaotic to the Gaussian density

$$\gamma(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2},$$

which in turn is the unique equilibrium of the Boltzmann-Kac equation.

When working with chaotic sequences on Kac's sphere, it is very desirable to have explicit rates of chaoticity (in N), and one natural way to quantify chaos is in the $L^2(d\sigma^N)$ sense. However, as explained in [3], the L^2 norm of a chaotic sequence tends to behave badly: it can grow exponentially with N . A better alternative is to use *entropy*: the *relative entropy* of $F^N \in \mathcal{P}(\mathbb{R}^N)$ with respect to $G^N(\mathbb{R}^N)$ is given by

$$H(F^N | G^N) = \int_{\mathbb{R}^N} h \log h dG^N = \int_{\mathbb{R}^N} \log h dF^N \geq 0, \quad h = \frac{dF^N}{dG^N},$$

and $H(F^N | G^N) = +\infty$ when F^N is not absolutely continuous with respect to G^N . A crucial advantage of the relative entropy over the L^2 norm is *extensivity*: for $f, g \in \mathcal{P}(\mathbb{R})$ such that $H(f | g) < \infty$, one has

$$H(f^{\otimes N} | g^{\otimes N}) = NH(f | g).$$

For a sequence $F^N \in \mathcal{P}(\mathcal{K}^N)$ which is f -chaotic, one expects a similar relation to hold approximately. The precise definition, introduced in [3, Definition 8], is the following:

Definition 2 (entropic chaos). For each $N \in \mathbb{N}$, let $F^N \in \mathcal{P}_{\text{sym}}(\mathcal{K}^N)$, and let $f \in \mathcal{P}(\mathbb{R})$ be such that $H(f | \gamma) < \infty$. The sequence $(F^N)_{N \in \mathbb{N}}$ is said to be entropically chaotic to f if it is Kac chaotic to f , and

$$\lim_{N \rightarrow \infty} \frac{1}{N} H(F^N | \sigma^N) = H(f | \gamma). \tag{1.1}$$

This definition can be thought of as “asymptotic extensivity” of the entropy. It is a stronger notion of chaos, involving all the N variables and not just a fixed number of marginals.

1.2 Measures rescaled to Kac's sphere

Given $f \in \mathcal{P}(\mathbb{R})$, consider the problem of finding a sequence F^N of measures supported on Kac's sphere that is f -chaotic, or better, f -entropically chaotic. One way of obtaining such a sequence is to take the tensor product $f^{\otimes N}$ and *condition* (restrict) it to Kac's sphere. The idea is that, if one assumes $\int_{\mathbb{R}} x^2 f(dx) = 1$, then, by the law of large numbers, one has $\sum_i x_i^2 \approx N$ under the law $f^{\otimes N}$; this means that $f^{\otimes N}$ is already concentrated around \mathcal{K}^N , and the conditioning should not change it too much. For a bounded density with finite fourth moment, it is proven in [3, Theorem 9] that the sequence obtained with this construction is indeed entropically chaotic to f . See also [4] for similar results in the context of *Boltzmann's sphere*.

In the present article, we propose the following alternative way of constructing a chaotic sequence on Kac's sphere, based on *rescaling* instead of conditioning:

Definition 3 (rescaled measure). For $x \in \mathbb{R}^N \setminus \{0\}$, we denote $\hat{x} \in \mathcal{K}^N$ the vector x rescaled to Kac's sphere, that is,

$$\hat{x} = \frac{\sqrt{N}}{|x|}x.$$

Similarly, for $F^N \in \mathcal{P}(\mathbb{R}^N)$ without an atom at the origin, we define the rescaled probability measure $\hat{F}^N \in \mathcal{P}(\mathcal{K}^N)$ as the push forward of F^N by the mapping $x \in \mathbb{R}^N \setminus \{0\} \mapsto \hat{x} \in \mathcal{K}^N$, that is,

$$\int_{\mathcal{K}^N} \phi(x) \hat{F}^N(dx) = \int_{\mathbb{R}^N} \phi(\hat{x}) F^N(dx).$$

Moreover, given $f \in \mathcal{P}(\mathbb{R})$ without an atom at 0, we denote $\hat{f}^N = \hat{F}^N$ for $F^N = f^{\otimes N}$.

For example, it can be easily seen that $\hat{\gamma}^N = \sigma^N$, by rotational symmetry of $\gamma^{\otimes N}$. This will be used several times throughout this article. Moreover, if $Z^N \sim \gamma^{\otimes N}$, then $|Z^N|$ and $\hat{Z}^N \sim \sigma^N$ are independent.

We remark that, to the best of our knowledge, the earliest use of this kind of rescaling in the setting of kinetic theory can be found in the proof of [9, Lemma 25], where the authors provide explicit rates of chaoticity for the uniform distribution on Boltzmann's sphere towards the Gaussian distribution on \mathbb{R}^3 . The first result for general rescaled measures seems to be found in [7, Section 5]; see also [6, 8]. In all of these references, some chaos estimates are proven using the 2-Wasserstein distance. The main goal of the present article is to investigate other notions of chaos satisfied by sequences of rescaled measures, in particular entropic chaos.

The main feature of the rescaled measure is its *simplicity*. For instance, its definition is straightforward and is valid for any $F^N \in \mathcal{P}(\mathbb{R}^N)$ that does not have an atom at the origin¹, while the conditioning procedure is typically applied only to tensor product measures $f^{\otimes N}$ and requires some smoothness and integrability assumptions on f , see [3, Definition 7]. Similarly, as we shall see, the results one can obtain for rescaled measures require less assumptions, and the proofs tend to be simpler. For instance, some of our proofs make use of classical limit theorems such as the Law of Large Numbers and the Central Limit Theorem, while the analysis of conditioned measures requires a refined local version of the latter (see [3, Appendix A]). Another nice feature

¹If $f \in \mathcal{P}(\mathbb{R})$ has an atom at 0 (but it's not the Dirac mass at 0), one can still define \hat{f}^N in such a way that it is f -chaotic, see [8] for details.

is that it is straightforward to sample from \hat{f}^N : simply generate $X^N \sim f^{\otimes N}$ and then compute $\hat{X}^N = \sqrt{N}X^N/|X^N| \sim \hat{f}^N$; this is an efficient way to generate a suitable initial condition when simulating Kac's particle system. For all these reasons, we believe that the rescaling procedure is a natural and convenient way, alternative to the usual conditioning method, of producing chaotic sequences on Kac's sphere.

1.3 Main results

In this work we prove several results concerning chaoticity and related properties for sequences of measures constructed via rescaling to Kac's sphere (Definition 3). We now summarize our main results:

- In **Theorem 7** we prove that the sequence $(\hat{f}^N)_{N \in \mathbb{N}}$, of rescaled tensor product measures, is always f -chaotic, under no assumptions on f other than unit energy and having no atom at 0. The proof is straightforward, relying only on the classical law of large numbers.
- In **Theorem 10** we show that if f has $2 + \epsilon$ moments and belongs to a broad subspace of "almost Lipschitz functions" in $L^1(\mathbb{R})$, which we introduce in Definition 9, then the k -marginals of \hat{f}^N converge in $L^1(\mathbb{R}^k)$ to $f^{\otimes k}$. The convergence is of order $N^{-\eta}$, with η explicitly given in (3.5).
- In **Theorem 27** we prove that the sequence $(\hat{f}^N)_{N \in \mathbb{N}}$ is always f -entropically chaotic, requiring only that f have unit energy and finite entropy relative to γ . Perhaps surprisingly, the proof is quite straightforward: one can easily see that $H(\hat{f}^N | \sigma^N)$ is always smaller than $NH(f | \gamma)$ by writing their difference as the negative of some relative entropy (the other inequality in (1.1) follows from [3, Theorem 12]).
- In **Theorem 34** we show that the sequence $(\hat{f}^N)_{N \in \mathbb{N}}$ is Fisher-information chaotic to f under mild assumptions (see Definition 33).

1.4 Structure of the article

We split our presentation as follows.

In Section 2 we study Kac's chaos. In Theorem 5 we prove that the rescaled measure of any f -chaotic sequence (not necessarily a tensor product) is also f -chaotic, under a uniform $p > 4$ moment assumption. In Lemma 6 we prove that the rescaled measure of a tensor product has exactly the same uniform (in N) moments as the the finite moments of the reference measure. In Theorem 7 we prove that the rescaling of a tensor product is always chaotic; we also recall [8, Theorem 3] for explicit rates of chaoticity in Wasserstein-2 distance, assuming additional moments. Finally, in Corollary 8 we provide similar estimates in the Wasserstein- p metric for $p > 2$.

Section 3 considers chaos in the sense of L^1 (see (3.1)). As mentioned above, we show in Theorem 10 that the k -marginals of \hat{f}^N converge in $L^1(\mathbb{R}^k)$ to $f^{\otimes k}$, under some conditions on f . In this section we also provide some examples and properties of the aforementioned "almost Lipschitz functions" in Lemmas 19 and 20. We end this section by comparing the result of Theorem 10 to a convergence result of the k -marginals of \hat{f}^N in the weak L^1 topology (i.e. against $L^\infty(\mathbb{R}^k)$ functions) which can be obtained from the Dunford and Pettis criterion whenever $H(f | \gamma) < \infty$. This section is independent of the rest.

Section 4 is devoted to entropic chaos. In Theorem 27 we prove that $(\hat{f}^N)_N$ is f -entropically chaotic, under minimal assumptions on f . In Theorem 22 we recall [3, Theorem 12] (giving one of the inequalities in (1.1)), for which we provide a simpler proof in the Appendix. In Theorem 32, we provide a quantitative rate of entropic chaos with mild polynomial order in N for the rescaled tensor product. The proof relies on

a control on the *Fisher information* of the sequence, provided in the next section, and on an improved version of [10, Theorem 4.17], given in Theorem 28, which relaxes the finite 6-moments condition on f to just $4 + \epsilon$. While proving this, we also improved [10, Theorem 4.13], which is a quantitative entropic chaos result concerning *conditioned tensor product measures* (defined in Definition 29), again by relaxing the moments requirements. This is done in Lemma 30.

In Section 5 we study the even stronger notion of Fisher information chaos (see Definition 33). We show in Theorem 34 that the rescaled tensor product measures $(\hat{f}^N)_N$ are Fisher-information chaotic to f under mild assumptions. The proof requires several computational results given by lemmas and propositions 35-39.

In Section 6 we give a general conclusion and mention some open problems. Finally, in the Appendix we provide proofs for some of the technical results.

1.5 Notation

Let us fix some notation:

- $\mathcal{P}(E)$ denotes the space of probability measures on the metric space E ; for $p \leq 1$, $\mathcal{P}_p(E)$ is the subspace of probability measures with finite p moment. $\mathcal{P}_{\text{sym}}(\mathbb{R}^N)$ denotes the space of symmetric probability measures on \mathbb{R}^N , i.e., those invariant under permutations of the N coordinates. $C_b(E)$ denotes the space of continuous and bounded real functions on E .
- Given $F^N \in \mathcal{P}(\mathbb{R}^N)$, we denote $\Pi_k F^N \in \mathcal{P}(\mathbb{R}^k)$ the projection of F^N on its first $k \leq N$ variables.
- $S^{N-1}(r) = \{x \in \mathbb{R}^N : |x|^2 = r\}$ is the $(N - 1)$ -dimensional sphere of radius r on \mathbb{R}^N , where $|x| = (x_1^2 + \dots + x_N^2)^{1/2}$ denotes the usual Euclidean norm. Thus, Kac's sphere for N particles is $\mathcal{K}^N = S^{N-1}(\sqrt{N})$.
- If $F^N \in \mathcal{P}(\mathbb{R}^N)$ has a density with respect to the Lebesgue measure, we will abuse notation and denote $F^N(x)$ its density at a point $x \in \mathbb{R}^N$.
- W_p denotes the p -Wasserstein distance on $\mathcal{P}_p(\mathbb{R}^N)$, that is, for $F^N, G^N \in \mathcal{P}_p(\mathbb{R}^N)$,

$$W_p(F^N, G^N) = \left(\inf_{\pi} \int_{\mathbb{R}^N \times \mathbb{R}^N} |x - y|^p \pi(dx, dy) \right)^{1/p},$$

where the infimum is taken over all $\pi \in \mathcal{P}(\mathbb{R}^N \times \mathbb{R}^N)$ with F^N and G^N as first and second marginals, respectively.

- ∇_S denotes the *spherical gradient* on the Kac's sphere \mathcal{K}^N : for any $F \in C^1(\mathcal{K}^N)$ and $y \in \mathcal{K}^N$,

$$\nabla_S F(y) = \nabla \tilde{F}(y) - \frac{1}{N} \nabla \tilde{F}(y) \cdot y,$$

where \tilde{F} is any C^1 extension of F to a neighbourhood of y in \mathbb{R}^N . $\nabla_S F(y)$ does not depend on the choice of \tilde{F} .

- The *relative Fisher information* of $F^N \in \mathcal{P}(\mathbb{R}^N)$ with respect to $G^N \in \mathcal{P}(\mathbb{R}^N)$ is given by

$$I(F^N | G^N) = \int_{\mathbb{R}^N} \frac{|\nabla h(x)|^2}{h(x)} G^N(dx), \quad h = \frac{dF^N}{dG^N},$$

and $I(F^N | G^N) = +\infty$ when F^N does not have a density with respect to G^N . The relative Fisher information of $F^N \in \mathcal{P}(\mathcal{K}^N)$ with respect to $G^N \in \mathcal{P}(\mathcal{K}^N)$ is defined similarly, but replacing the usual gradient ∇ by the spherical gradient ∇_S .

2 Kac's chaos

We start by stating a weak law of large numbers in L^2 for chaotic sequences. The proof involves a computation similar to the one in [15, Proposition 2.2].

Lemma 4 (law of large numbers for chaotic sequences). *Let $F^N \in \mathcal{P}_{\text{sym}}(\mathbb{R}^N)$ be an f -chaotic sequence, for some $f \in \mathcal{P}(\mathbb{R})$, such that $\sup_N \int_{\mathbb{R}^N} |x_1|^p F^N(dx) < \infty$ for some $p > 2$. Let $X^N = (X_1^N, \dots, X_N^N)$ be F^N -distributed. Then, for any continuous $\phi : \mathbb{R} \rightarrow \mathbb{R}$ with at most linear growth, we have*

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \phi(X_i^N) = \int_{\mathbb{R}} \phi(x) f(dx), \quad \text{in } L^2.$$

Proof. For any such ϕ , it is clear that $\phi(X_1^N)$ and $\phi(X_1^N)\phi(X_2^N)$ are uniformly (in N) integrable: by the uniform bound of the p -th moment and the linear growth assumption of ϕ , it is easily seen that $\sup_N \mathbb{E}[\phi(X_1^N)^p] < \infty$ and $\sup_N \mathbb{E}[(\phi(X_1^N)\phi(X_2^N))^{p/2}] < \infty$. Consequently, using f -chaoticity, we get

$$\lim_{N \rightarrow \infty} \mathbb{E}[\phi(X_1^N)] = m, \quad \text{and} \quad \lim_{N \rightarrow \infty} \mathbb{E}[\phi(X_1^N)\phi(X_2^N)] = m^2, \quad (2.1)$$

where $m = \int_{\mathbb{R}} \phi(x) f(dx)$. Using exchangeability, we have:

$$\begin{aligned} & \mathbb{E} \left[\left(\frac{1}{N} \sum_{i=1}^N \phi(X_i^N) - m \right)^2 \right] \\ &= \mathbb{E} \left[\frac{1}{N^2} \sum_{i=1}^N \phi(X_i^N)^2 + \frac{1}{N^2} \sum_{i \neq j} \phi(X_i^N)\phi(X_j^N) - \frac{2m}{N} \sum_{i=1}^N \phi(X_i^N) + m^2 \right] \\ &= \frac{\mathbb{E}[\phi(X_1^N)^2]}{N} + \frac{N-1}{N} \mathbb{E}[\phi(X_1^N)\phi(X_2^N)] - 2m\mathbb{E}[\phi(X_1^N)] + m^2, \end{aligned}$$

which converges to 0 as $N \rightarrow \infty$ thanks to (2.1). □

The next theorem shows that the rescaled measures of an f -chaotic sequence are also f -chaotic, under a mild assumption of bounded p -moments for some $p > 4$. The proof is straightforward, and relies on the law of large numbers stated in Lemma 4. This showcases the simplicity of the rescaled measures.

Theorem 5 (chaos for rescaled measures). *Let $(F^N)_{N \in \mathbb{N}}$ be an f -chaotic sequence, for some $f \in \mathcal{P}(\mathbb{R})$ satisfying $\int_{\mathbb{R}} x^2 f(dx) = 1$. Assume that F^N does not have an atom at the origin, and that $\sup_N \int_{\mathbb{R}^N} |x_1|^p F^N(dx) < \infty$ for some $p > 4$. Then \hat{F}^N is f -chaotic.*

Proof. Let $X^N = (X_1^N, \dots, X_N^N)$ be F^N -distributed, and recall that $\hat{X}^N = (\hat{X}_1^N, \dots, \hat{X}_N^N)$ denotes the vector X^N rescaled to Kac's sphere, see Definition 3. We need to show that for any $\phi : \mathbb{R}^k \rightarrow \mathbb{R}$ L -Lipschitz continuous and bounded one has

$$\lim_{N \rightarrow \infty} \mathbb{E}[\phi(\hat{X}_1^N, \dots, \hat{X}_k^N)] = \int_{\mathbb{R}^k} \phi(x_1, \dots, x_k) f(dx_1) \cdots f(dx_k). \quad (2.2)$$

Without loss of generality, we assume that $E_N := \int_{\mathbb{R}^N} x_1^2 F^N(dx) = 1$ for all N (if not, we can easily reduce to this case by working with $X^N/\sqrt{E_N}$, which is also f chaotic since $\lim_N E_N = 1$ thanks to the f -chaoticity and the bounded p -moment assumptions). Call $Q_N = \frac{1}{N} \sum_i (X_i^N)^2$ and $Y = (X_1^N, \dots, X_k^N)$, thus $\hat{X}_i^N = Y_i/\sqrt{Q_N}$ for $i = 1, \dots, k$. Since

(F^N) is f -chaotic, we know that $\mathbb{E}[\phi(Y)]$ converges to the r.h.s. of (2.2); thus, it suffices to show that the following vanishes as $N \rightarrow \infty$:

$$\begin{aligned} \left| \mathbb{E}[\phi(Q_N^{-1/2}Y)] - \mathbb{E}[\phi(Y)] \right| &\leq L\mathbb{E} \left[\left| (Q_N^{-1/2} - 1)Y \right| \right] \\ &\leq kL\mathbb{E} \left[\left| Q_N^{-1/2} - 1 \right| \frac{1}{N} \sum_{i=1}^N |X_i^N| \right] \\ &\leq kL\mathbb{E} \left[\left| Q_N^{-1/2} - 1 \right| Q_N^{1/2} \right] \\ &= kL\mathbb{E} \left[\left| Q_N^{1/2} - 1 \right| \right] \\ &\leq kL\mathbb{E} [|Q_N - 1|], \end{aligned} \tag{2.3}$$

where we have used exchangeability. Finally, thanks to the uniform (in N) bound on the moment of F^N of order $p > 4$, we can apply Lemma 4 to the sequence of distributions $(\text{Law}((X_1^N)^2, \dots, (X_N^N)^2))_{N \in \mathbb{N}}$ with $\phi(x) = x$, and deduce that $\lim_N Q_N = 1$ in L^2 ; thus, (2.3) converges to 0 as $N \rightarrow \infty$. \square

We now turn our attention to rescaled tensor products $\hat{f}^N := \hat{F}^N$ with $F^N = f^{\otimes N}$, for some $f \in \mathcal{P}(\mathbb{R})$. The next lemma shows that \hat{f}^N has the same uniformly (in N) bounded one-particle moments as the finite moments of the original measure f :

Lemma 6 (moments for rescaled tensor products). *Let $f \in \mathcal{P}_2(\mathbb{R})$, without an atom at 0. Then, for any $p \geq 2$,*

$$\sup_N \int_{\mathbb{R}^N} |x_1|^p \hat{f}^N(dx) < \infty \quad \text{if and only if} \quad \int_{\mathbb{R}} |x|^p f(dx) < \infty.$$

Proof. We first prove the direct implication. Let $X^N = (X_1, \dots, X_N)$ be a collection of i.i.d. and f -distributed random variables, thus $\hat{X}^N \sim \hat{f}^N$. Call $Q_N := \frac{1}{N} \sum_i X_i^2$, thus $\hat{X}_i^N = X_i/\sqrt{Q_N}$ for all $i = 1, \dots, N$. Denoting $E := \int_{\mathbb{R}} x^2 f(dx)$, we have:

$$\begin{aligned} \mathbb{E}[|\hat{X}_N^N|^p] &= \mathbb{E} \left[|X_N|^p \left(\frac{X_N^2}{N} + \frac{N-1}{N} Q_{N-1} \right)^{-p/2} \right] \\ &= \left(\frac{N}{N-1} \right)^{p/2} \int_{\mathbb{R}} |x|^p \mathbb{E} \left[\left(\frac{x^2}{N-1} + Q_{N-1} \right)^{-p/2} \right] f(dx) \\ &\geq \left(\frac{N}{N-1} \right)^{p/2} \int_{\mathbb{R}} |x|^p \left(\frac{x^2}{N-1} + E \right)^{-p/2} f(dx), \end{aligned} \tag{2.4}$$

where in the last step we used Jensen's inequality. By the monotone convergence theorem, we deduce that

$$\liminf_{N \rightarrow \infty} \mathbb{E}[|\hat{X}_N^N|^p] \geq E^{-p/2} \int_{\mathbb{R}} |x|^p f(dx),$$

which proves the direct implication.

Now we prove the converse: given any $A > 0$, we split the expectation in (2.4) in the cases $Q_{N-1} \geq A$ and $Q_{N-1} < A$, which gives

$$\begin{aligned} \mathbb{E}[|\hat{X}_N^N|^p] &\leq \left(\frac{N}{N-1} \right)^{p/2} \int_{\mathbb{R}} |x|^p \left\{ A^{-p/2} + \left(\frac{x^2}{N-1} \right)^{-p/2} \mathbb{P}(Q_{N-1} < A) \right\} f(dx) \\ &= \left(\frac{N}{N-1} \right)^{p/2} A^{-p/2} \int_{\mathbb{R}} |x|^p f(dx) + N^{p/2} \mathbb{P}(Q_{N-1} < A). \end{aligned}$$

Thus, it suffices to show that for some $A > 0$, the second term in the last expression is bounded uniformly in N . To this end, for each $i = 1, \dots, N - 1$, consider the random variable $Y_i := \mathbf{1}_{\{X_i^2 \geq E\}}$, thus Y_1, \dots, Y_{N-1} are i.i.d. Bernoulli(q) for $q = \mathbb{P}(X_1^2 \geq E)$, which is strictly positive because $\mathbb{E}[X_1^2] = E$. Clearly $EY_i \leq X_i^2$, thus for $A = Eq/2$ we have

$$\mathbb{P}(Q_{N-1} < A) \leq \mathbb{P}\left(\frac{1}{N-1} \sum_{i=1}^{N-1} Y_i < \frac{q}{2}\right) \leq e^{-q(N-1)/8},$$

where we have used the Chernoff-type bound $\mathbb{P}(\sum_{i=1}^n Y_i \leq n(1 - \delta)q) \leq e^{-\delta^2 nq/2}$ (valid for any $0 < \delta < 1$), with $\delta = 1/2$. We thus deduce that $\sup_N N^{p/2} \mathbb{P}(Q_{N-1} < A) < \infty$ for $A = Eq/2$, as desired. This concludes the proof. \square

The next theorem shows that \hat{f}^N is always f -chaotic, under no assumption other than unit energy and no atom at 0. The proof is straightforward, and relies on the usual law of large numbers. Moreover, under the additional assumption that f has finite p -moment for some $p > 2$, using [8, Theorem 3], we can provide an explicit rate of chaoticity of polynomial order, in the 2-Wasserstein distance. We remark that this is a “strong” chaos result (see also Corollary 8), in the sense that it compares \hat{f}^N and $f^{\otimes N}$ directly, and not just a fixed number of marginals.

Theorem 7 (chaos for rescaled tensor products). *Let $f \in \mathcal{P}(\mathbb{R})$, without an atom at 0, and such that $\int_{\mathbb{R}} x^2 f(dx) = 1$. Then the sequence $(\hat{f}^N)_N$ is chaotic to f . Moreover, if $f \in \mathcal{P}_p(\mathbb{R})$ for some $p > 2$, then we have the following quantitative rate: there exists a constant C depending only on p and $\int_{\mathbb{R}} |x|^p f(dx)$, such that*

$$\frac{1}{N} W_2(\hat{f}^N, f^{\otimes N})^2 \leq C \begin{cases} N^{-1/2}, & p > 4, \\ N^{-1/2} \log(N), & p = 4, \\ N^{-(1-2/p)}, & p \in (2, 4). \end{cases} \tag{2.5}$$

Proof. Using the same argument and notation as in the proof of Theorem 5, the f -chaoticity of \hat{f}^N , assuming only that $\int_{\mathbb{R}} x^2 f(dx) = 1$, follows from (2.3) and the usual law of large numbers in L^1 for i.i.d. sequences. The proof of (2.5) can be found in [8, Theorem 3], so we omit it here. \square

As a corollary, we can bound $W_r(\hat{f}^N, f^{\otimes N})^r$ for $r > 2$ by imposing finite $p > r$ moments on f (see [10, Equation 2.11] for a special case).

Corollary 8. *Let $f \in \mathcal{P}_p(\mathbb{R})$ for some $p > 2$, without an atom at 0. Then, for any $2 < r < p$, there exists a constant $C > 0$ such that, for $b = \frac{p-r}{p-2} < 1$, we have*

$$\frac{1}{N} W_r(\hat{f}^N, f^{\otimes N})^r \leq C \begin{cases} N^{-b/2}, & p > 4, \\ N^{-b/2} \log(N)^b, & p = 4, \\ N^{-b(1-2/p)}, & p \in (2, 4). \end{cases}$$

Proof. Let U^N and V^N be random vectors on \mathbb{R}^N such that $U^N \sim f^{\otimes N}$, $V^N \sim \hat{f}^N$, and $W_2(f^{\otimes N}, \hat{f}^N)^2 = \mathbb{E}|U^N - V^N|^2$. Moreover, our hypotheses, together with Lemma 6, imply that $\sup_N (\mathbb{E}|U_1^N|^p + \mathbb{E}|V_1^N|^p) =: M_p < \infty$. Noting that $p = \frac{r-2b}{1-b}$ and using Hölder's

inequality with exponents $\frac{1}{1-b}$ and $\frac{1}{b}$, we have:

$$\begin{aligned} \frac{1}{N} \mathbb{E}|U^N - V^N|^r &= \frac{1}{N} \mathbb{E}|U^N - V^N|^{r-2b} |U^N - V^N|^{2b} \\ &\leq \frac{1}{N} (\mathbb{E}|U^N - V^N|^p)^{1-b} (\mathbb{E}|U^N - V^N|^2)^b \\ &= \left(\frac{1}{N} \mathbb{E}|U^N - V^N|^p \right)^{1-b} \left(\frac{1}{N} \mathbb{E}|U^N - V^N|^2 \right)^b \\ &\leq (2^{p-1} M_p)^{1-b} \left(\frac{1}{N} W_2(f^{\otimes N}, \hat{f}^N)^2 \right)^b, \end{aligned}$$

where we used exchangeability and the inequality $|U_i^N - V_i^N|^p \leq 2^{p-1}(|U_i^N|^p + |V_i^N|^p)$. Since $\mathbb{E}|U^N - V^N|^r$ is an upper estimate of $W_r(f^{\otimes N}, \hat{f}^N)^r$, the conclusion follows from Theorem 7. \square

3 Chaos in L^1

In this section we study chaoticity of $\{\hat{f}^N\}$ to f in the sense of L^1 . More specifically, chaoticity in L^1 means that

$$\lim_{N \rightarrow \infty} \|\Pi_k \hat{f}^N - f^{\otimes k}\|_{L^1(\mathbb{R}^k)} = 0, \quad \forall k \geq 1, \tag{3.1}$$

where $\Pi_k \hat{f}^N \in \mathcal{P}(\mathbb{R}^k)$ denotes the first k -th marginal of \hat{f}^N . We show in Theorem 10 that (3.1) holds (with quantitative rates) for all f in a large subset of $L^1(\mathbb{R})$, which we call “almost Lipschitz functions”, defined formally in Definition 9.

The proof of Theorem 10 is somewhat technical, so we now describe the main ideas. Fix $f \in L^1(\mathbb{R})$ with $\int_{\mathbb{R}} x^2 f(x) dx = 1$ and consider the action of $\Pi_k \hat{f}^N$ when integrated against a function $\phi \in C_0^1(\mathbb{R}^k)$. Let $L_N[\phi] = \int_{\mathbb{R}^k} \Pi_k \hat{f}^N(x) \phi(x) dx$. Then

$$L_N[\phi] = \int_{\mathcal{K}^N} \phi(x_1, \dots, x_k) \frac{d\hat{f}^N}{d\sigma^N}(x) \sigma^N(dx) = \int_{\mathbb{R}^N \setminus \{0\}} f^{\otimes N}(x) \phi \left(\frac{\sqrt{N}}{|x|} (x_1, \dots, x_k) \right) d^N x.$$

We decompose \mathbb{R}^N into $\mathbb{R}^k \times \mathbb{R}^{N-k}$, $x = (y, z)$, and note that in the last integral above z only appears through $|z|^2$. Hence, if we introduce

$$\mu_{N-k} = \text{law} (X_1^2 + \dots + X_{N-k}^2), \quad (X_1, \dots, X_{N-k}) \sim f^{\otimes(N-k)} \tag{3.2}$$

then

$$L_N[\phi] = \int_{s \in (0, \infty)} \mu_{N-k}(ds) \int_{y \in \mathbb{R}^k} f^{\otimes k}(y) \phi \left(\sqrt{\frac{N}{s + |y|^2}} y \right) d^k y. \tag{3.3}$$

We want to “remove” the expression $\sqrt{\frac{N}{s + |y|^2}}$ above. By a law of large numbers, we expect s to be concentrated around $N - k$. In Lemma 14 (and Corollary 15), we show that, with a cost that vanishes as $N \rightarrow \infty$ we can restrict s to an interval of the form $[N - N^\alpha, N + N^\alpha]$. For this step we require f to have a finite moment of order greater than 2. Next, when s lies inside this interval, we would like $|y|^2$ to be negligible compared to s . Thus, we show that we can replace the mapping $y \mapsto \sqrt{\frac{N}{s + |y|^2}} y$ by a mapping of the form $y \mapsto \psi(y) = \sqrt{\frac{N}{s + \min\{|y|^2, N^{\alpha/2}\}}} y$. This mapping is invertible and close to the identity mapping. In Lemma 11 we compute its inverse and the determinant of its Jacobian explicitly, and in Lemma 12 we quantify its distance from the identity map. Chaoticity in L^1 (recall (3.1)) will be proven once we show that the quantity

$$\|F - F(\psi^{-1}(x))\| \det(D\psi^{-1}(x))\|_{L^1(\mathbb{R}^k)} \tag{3.4}$$

is small when ψ is close to the identity. Here $F = f^{\otimes k}$. It is here that we need f to be more than just an element of $L^1(\mathbb{R})$. We first obtain an upper bound to (3.4) in Lemma 16 for f Lipschitz, and relax this condition to “almost Lipschitz” in Lemma 17. For ease, all the lemmas mentioned above are stated after the proof of Theorem 10.

We now specify what we mean by “almost Lipschitz” functions. We recall that any function in $L^1(\mathbb{R}^k)$ can be approximated in $L^1(\mathbb{R}^k)$ by Lipschitz functions (e.g. consider the convolutions of f with $\epsilon^{-\frac{k}{2}} \exp(-\pi|x|^2/\epsilon)$). The almost Lipschitz functions, defined next, ask for a quantitative bound on the Lipschitz constant of the function g_ϵ that ϵ -approximates f in $L^1(\mathbb{R})$.

Definition 9 (The almost Lipschitz spaces). *Let $r \geq 0$ and $L < \infty$. We first define the set $\text{ALip}(r, L)$ as follows*

$$\text{ALip}(r, L) = \left\{ f \in L^1(\mathbb{R}) : \forall \epsilon > 0, \exists g_\epsilon \in L^1(\mathbb{R}) \text{ and Lipschitz, such that } \|f - g_\epsilon\|_{L^1(\mathbb{R})} \leq \epsilon \text{ and } \|g_\epsilon\|_{\text{Lip}} \leq L\epsilon^{-r} \right\}.$$

The almost-Lipschitz space is given by

$$\text{ALip}(r) = \bigcup_{L < \infty} \text{ALip}(r, L).$$

As we will show in Lemma 19, after the proof of Theorem 10, $\text{ALip}(r)$ contains all piecewise Lipschitz functions that have at most countably many jumps and countably many blowups of the type $|x - x_0|^{-a}$ with $a \leq a_0$, for some fixed $a_0 = a_0(r) < 1$. We also in Lemma 20 that $\text{ALip}(r)$ is a subset of a Besov space.

We are now ready to state and prove Theorem 10.

Theorem 10. *Fix $r \geq 0$ and let f be a probability density on \mathbb{R} that belongs to $\text{ALip}(r)$. Assume that $\int_{\mathbb{R}} x^2 f(x) dx = 1$ and that $\int_{\mathbb{R}} |x|^{2+\delta} f(x) dx < \infty$ for some $\delta \in (0, 2]$. Then there exists a constant C_0 independent of N such that, for all $k \leq N$,*

$$\|\Pi_k \hat{f}^N - f^{\otimes k}\|_{L^1(\mathbb{R}^k)} \leq C_0 N^{-\eta},$$

where η is given by

$$\eta = \frac{\delta}{k + 5 + \delta + r(2 + \delta)}. \tag{3.5}$$

Proof. Without loss of generality, we will take N to be at least N_0 given by

$$N_0(k, \delta) = (2k)^{1+\delta/2}. \tag{3.6}$$

Let μ_{N-k} and $L_N[\phi]$ be as in above (see equations (3.2) and (3.3)).

We fix $q \in (0, 1)$, which will be optimized at the end of the proof. From (3.3), we obtain:

$$L_N[\phi] = E_1[\phi] + \int_{|s-N| \leq N^{1-q}} \mu_{N-k}(ds) \left(\int_{\mathbb{R}^k} f^{\otimes k}(y) \phi \left(\sqrt{\frac{N}{|y|^2 + s}} y \right) d^k y \right)$$

where $E_1[\phi]$ is the integral above when $|s - N| > N^{1-q}$. Clearly, $|E_1[\phi]| \leq \|\phi\|_{L^\infty} \mathbb{P}(\tilde{S}_{N,k})$, with

$$\tilde{S}_{N,k} = \{|X_1^2 + \dots + X_{N-k}^2 - N| > N^{1-q}\}.$$

We next find the rates at which $\mathbb{P}(\tilde{S}_{N,k})$ (and hence $E_1[\phi]$) approaches zero. A quantitative law of large numbers from [17], reproduced in Lemma 14 below, implies that (details in Corollary 15)

$$\mathbb{P}(\tilde{S}_{N,k}) \leq 2^{2+\delta/2} N^{-(\delta/2 - (1+\delta/2)q)} \mathbb{E}|X_1|^{2+\delta}.$$

This makes

$$|E_1[\phi]| \leq 8\|\Phi\|_\infty \mathbb{E}|X_1|^{2+\delta} N^{-(\delta/2-(1+\delta/2)q)}. \tag{3.7}$$

Next, we analyze further $L_N[\phi]$ and cutoff $|y|^2$ at $(N - k)^{(1-q)/2}$ so that we can write

$$L_N[\phi] = E_1[\phi] + E_2[\phi] + \int_{\{|s-N| \leq N^{1-q}\}} \mu_{N-k}(ds) \left(\int_{\mathbb{R}^k} f^{\otimes k}(y) \phi(\psi_{N,s}(y)) d^k y \right), \tag{3.8}$$

where $\psi_{N,s}$ is defined by

$$\psi_{N,s}(y) = \sqrt{\frac{N}{\min\{|y|^2, (N - k)^{(1-q)/2}\} + s}} y.$$

The term E_2 can be bounded. Introducing the notation T_R for the tail $\int_{B(0,R)^c} f^{\otimes k}$, we have: $|E_2[\phi]| \leq 2\|\phi\|_{L^\infty} T_{N^{(1-q)/4}}$. The finite $(2 + \delta)$ -moment assumption on f , allows us to obtain a bound for the tail term $T_{N^{(1-q)/4}}$.

We observe that whenever $x_1^2 + \dots + x_k^2 \geq R^2$, we have

$$|x_1|^{2+\delta} + \dots + |x_k|^{2+\delta} \geq k \frac{R^{2+\delta}}{k^{1+\delta/2}}.$$

Thus, we have:

$$T_R \leq \int_{\mathbb{R}^k} \frac{k^{1+\delta/2}}{kR^{2+\delta}} \sum_{j=1}^k |x_j|^{2+\delta} f^{\otimes k}(x) d^k x = \frac{k^{1+\delta/2}}{R^{2+\delta}} \mathbb{E}|X_1|^{2+\delta},$$

which implies the following.

$$|E_2[\phi]| \leq 2\|\phi\|_{L^\infty} k^{1+\delta/2} \mathbb{E}|X_1|^{2+\delta} N^{-(1-q)(2+\delta)/4}. \tag{3.9}$$

We treat next the y integral term in (3.8) and show that it approximately equals $\int_{\mathbb{R}^k} f^{\otimes k}(y) \phi(y) d^k y$ uniformly over the integration range of s . In Lemmas 11 and 12, we show that $\psi_{N,s}^{-1}(y) \approx y$ in the sense that

$$|\psi_{N,s}^{-1}(y) - y| \leq \epsilon_N |y|, \quad \|D\psi_{N,s}^{-1}(y) - \text{Id}\| \leq \epsilon_N$$

with $\epsilon_N \leq C(k, q)N^{-q}$. Thus, we can use Lemma 17 below with $\beta = 2 + \delta$ to obtain

$$\|f^{\otimes k}(x) - f^{\otimes k}(\psi_{N,s}^{-1}(x))\|_{L^1(\mathbb{R}^k)} \leq CN^{-q} \frac{2+\delta}{k+3+\delta}. \tag{3.10}$$

which holds uniformly over the range $s \in [N - N^{1-q}, N + N^{1-q}]$. It is here that we use the assumption (3.6) which implies that $N \geq (2k)^{1/(1-q)}$. Here we have also substituted ϵ_N by CN^{-q} . Finally, since $\mu_{N-k}([N - N^{1-q}, N + N^{1-q}]) < 1$, the s -integral in (3.8) comes with this factor < 1 . Thus, we can instead compare $L_N[\phi]$ to $\mu_{N-k}([N - N^{1-q}, N + N^{1-q}]) \int_{\mathbb{R}^k} f^{\otimes k}(x) \phi(x) d^k x$. This "cost" is comparable to our bound on $|E_1[\phi]|$.

Combining the bounds in (3.7), (3.9), and (3.10), and the last observation concerning the factor $\mu_{N-k}([N - N^{1-q}, N + N^{1-q}]) < 1$, we obtain the following bound.

$$\begin{aligned} \|\Pi_k \hat{f}^N - f^{\otimes k}\|_{L^1(\mathbb{R}^k)} &\leq C(k, \delta, L, \mathbb{E}|X_1|^{2+\delta}) N^{-q(2+\delta)/(k+3+\delta)} + \\ &\quad C(\delta, \mathbb{E}|X_1|^{2+\delta}) N^{-(\delta/2-q(1+\delta/2))} + \\ &\quad C(k, \mathbb{E}|X_1|^{2+\delta}) N^{-(1-q)(2+\delta)/4}, \end{aligned}$$

whenever $N \geq \max\{2k, (2k\mathbb{E}X_1^2)^{1/(1-q)}\}$. This holds true for any $q \in (0, 1)$. It remains to choose q in order to maximize the slowest decay rate. This is done in Lemma 18, which implies that

$$\|\Pi_k \hat{f}^N - f^{\otimes k}\|_{L^1(\mathbb{R}^k)} \leq C(k, \delta, L, \mathbb{E}|X_1|^{2+\delta}) N^{-\delta/(k+5+\delta+r(2+\delta))}.$$

Here, the exponent $1 + \delta/2$ in (3.6) comes from a simple upper bound to $1/(1 - q)$ for the above optimal q . □

Lemma 11. Let a, b , and c be positive numbers, and consider the transformation $\psi : \mathbb{R}^k \rightarrow \mathbb{R}^k$ given by

$$z = \psi(x) = \left(\frac{a}{b + \min\{|x|^2, c\}} \right)^{1/2} x. \tag{3.11}$$

Then this transformation is one-to-one and onto; its inverse is given by

$$\psi^{-1}(x) = \begin{cases} \sqrt{\frac{b}{a-|z|^2}} z, & |z| \leq \sqrt{\frac{ac}{b+c}} \\ \sqrt{\frac{b+c}{a}} z, & |z| > \sqrt{\frac{ac}{b+c}} \end{cases}. \tag{3.12}$$

Also ψ^{-1} is differentiable away from the set $\{|z| = \sqrt{\frac{ac}{b+c}}\}$ with derivative:

$$D\psi^{-1}(z) = \begin{cases} \sqrt{\frac{b}{a-|z|^2}} \left(\text{Id} - \frac{zz^T}{a-|z|^2} \right), & |z| < \sqrt{\frac{ac}{b+c}} \\ \sqrt{\frac{b+c}{a}} \text{Id}, & |z| > \sqrt{\frac{ac}{b+c}} \end{cases}. \tag{3.13}$$

Finally, the Jacobian of ψ^{-1} is given as follows

$$|\det D\psi^{-1}(z)| = \begin{cases} \left(\frac{b}{a-|z|^2} \right)^{\frac{k}{2}} \left(1 - \frac{|z|^2}{a-|z|^2} \right), & |z| < \sqrt{\frac{ac}{b+c}} \\ \left(\frac{b+c}{a} \right)^{\frac{k}{2}}, & |z| > \sqrt{\frac{ac}{b+c}} \end{cases}$$

Proof. The proof of (3.11)–(3.12)–(3.13) follows from a direct computation. In order to compute the Jacobian, one can use the identity $\det(\text{Id} - a_0 vv^T) = 1 - a_0|v|^2$ for any $a_0 \in \mathbb{R}$ and $v \in \mathbb{R}^k$. \square

In the following lemma, we find an explicit number ϵ_N for which $|\psi^{-1}(z) - z| \leq \epsilon_N|z|$ and $\|D\psi^{-1}(z) - \text{Id}\| \leq \epsilon_N$, and $|1 - |\det D\psi^{-1}(x)|| \leq \epsilon_N$ whenever (A, B, C) depend on N and equal $(N, N + uN^{1-q}, N^{(1-q)/2})$ for some $q \in (0, 1)$ and $u \in (-1, 1)$.

Lemma 12. Fix $q \in (0, 1)$, and $u \in (-1, 1)$. Let ψ be given by (3.11) with $(a, b, c) = (a_N, b_N, c_N) = (N, N + uN^{1-q}, N^{(1-q)/2})$. Then there is a number $C(k, q) < \infty$ such that if $N \geq 2$ and $\epsilon_N = C(k, q)N^{-q}$, then

$$|\psi^{-1}(z) - z| \leq \epsilon_N|z|, \|D\psi^{-1}(z) - \text{Id}\| \leq \epsilon_N, |1 - |\det D\psi^{-1}(z)|| \leq \epsilon_N$$

Taking $C(k, q) = \frac{10+2k(1+2k/(1-2^{-q})^{k/2})}{1-2^{-q}}$ suffices.

Remark 13. In fact, it can be shown that

$$\begin{aligned} |\psi^{-1}(z) - z| &\leq \frac{4}{1-2^{-q}} N^{-q} |z| \\ \|D\psi^{-1}(z) - \text{Id}\| &\leq \left(4 \frac{(1+2^{-q})}{1-2^{-q}} + 1 \right) N^{-q} \\ |1 - |\det D\psi^{-1}(z)|| &\leq \frac{1+2k \left(1 + \frac{4N^{-q}}{1-2^{-q}} \right)^{k/2}}{1-2^{-q}} N^{-q} \end{aligned}$$

Proof of Lemma 12. The proof follows by studying how the coefficients of z in (3.12) and of the matrices in (3.13) compare to 1. It boils down to showing the following three identities which can be proven using elementary computations. In the following $z_0^2 = a_N c_N / (b_N + c_N)$ as hinted at in (3.12).

- $\left| \frac{b_N+c_N}{a_N} - 1 \right| \leq 2N^{-q}$, $N \geq 1$ (Thus, $\left| \sqrt{\frac{b_N+c_N}{a_N}} - 1 \right| \leq 2N^{-q}$
and $\left| \left(\frac{b_N+c_N}{a_N} \right)^l - 1 \right| \leq 2lN^{-q}(1 + 2N^{-q})^{l-1}$, for all $l > 1$)
- $\left| \frac{b_N}{a_N-|z|^2} - 1 \right| \leq \frac{4}{1-2^{-q}}N^{-q}$, whenever $z^2 \leq z_0^2$
and $N \geq 2$. (Thus $\left| \sqrt{\frac{b_N}{a_N-|z|^2}} - 1 \right| \leq \frac{4}{1-2^{-q}}N^{-q}$ and
 $\left| \left(\frac{b_N}{a_N-|z|^2} \right)^l - 1 \right| \leq \frac{4}{1-2^{-q}}lN^{-q} (1 + 4N^{-q}/(1 - 2^{-q}))^{l-1}$ for all $l > 1$).
- $\frac{z^2}{a_N-z^2} \leq \frac{N^{-(1-q)/2}}{1-2^{-q}}$, whenever $z^2 \leq z_0^2$, and $N \geq 2$. □

We now state without proof a result of von Bahr and Esseen [17].

Lemma 14 (A Quantitative Law of Large Numbers). *Let Y_1, \dots, Y_N be i.i.d. random variables with $\mathbb{E}Y_1 = 0$ and $\mathbb{E}|Y_1|^s < \infty$, where $1 \leq s \leq 2$. Then*

$$\mathbb{E} [N^{-s}|Y_1 + \dots + Y_N|^s] \leq 2N^{-(s-1)}\mathbb{E}|Y_1|^s.$$

Corollary 15. *Let X_1, \dots, X_N ($N > k$) be i.i.d., $\mathbb{E}|X_1|^{2+\delta} < \infty$, and let $q \in (0, 1)$. Define $\tilde{S}_{N,k}$ as follows:*

$$\tilde{S}_{N,k} = \{ |X_1^2 + \dots + X_{N-k}^2 - N\mathbb{E}X_1^2| > N^{1-q} \}.$$

Then, whenever $N \geq \max \{ 2k, (2k\mathbb{E}X_1^2)^{1/(1-q)} \}$, we have:

$$\mathbb{P}(\tilde{S}_{N,k}) \leq 16N^{-(\bar{\delta}/2-(1+\bar{\delta}/2)q)}\mathbb{E}|X_1|^{2+\bar{\delta}}. \tag{3.14}$$

Here $\bar{\delta} = \min\{2, \delta\}$.

Proof. Without loss of generality, we can let $\delta = \min\{\delta, 2\}$. Let the Y_j be as in Lemma 14. And let M_j be given by

$$M_j = \frac{Y_1 + \dots + Y_j}{j}.$$

It follows from Tchebyshev's inequality and Lemma 14 that for any $\beta_N > 0$, we have

$$\mathbb{P}(|M_N| > \beta_N) \leq 2N^{1-s}\beta_N^{-s}\mathbb{E}|Y_1|^s. \tag{3.15}$$

We will let $Y_j = X_j^2 - \mathbb{E}X_j^2$ and $s = (2 + \delta)/2$. Thus,

$$\mathbb{E}|Y_1|^s \leq 2^{1+\delta/2}\mathbb{E}|X_1|^{2+\delta}.$$

We now note that

$$\begin{aligned} \mathbb{P}(\tilde{S}_{N,k}^c) &= \mathbb{P}\left(\left| M_{N-k} - \frac{k}{N-k}\mathbb{E}X_1^2 \right| < \frac{N^{1-q}}{N-k} \right) \\ &= \mathbb{P}\left(\frac{k}{N-k}\mathbb{E}X_1^2 - \frac{N^{1-q}}{N-k} < M_{N-k} < \frac{k}{N-k}\mathbb{E}X_1^2 + \frac{N^{1-q}}{N-k} \right) \\ &\geq \mathbb{P}\left(\frac{\frac{1}{2}N^{1-q}}{N-k} - \frac{N^{1-q}}{N-k} < M_{N-k} < \frac{N^{1-q}}{N-k} \right) \\ &\geq \mathbb{P}\left(|M_{N-k}| < \frac{1}{2} \frac{N^{1-q}}{N-k} \right) \geq \mathbb{P}\left(|M_{N-k}| < \frac{1}{2}N^{-q} \right). \end{aligned}$$

In the last two steps above we used the assumptions $N \geq (2k\mathbb{E}X_1^2)^{1/(1-q)}$ and $N \geq 2k$. Using (3.15) with $s = 1 + \delta/2$ we obtain:

$$\mathbb{P}(\tilde{S}_{N,k}) \leq 2^{3+\delta/2}N^{-(\delta/2-(1+\delta/2)q)}\mathbb{E}|X_1|^{2+\delta},$$

which gives (3.14). □

The following two Lemmas are key to proving Theorem 10. Lemma 16 shows in a quantitative way that for a Lipschitz, L^1 function g and a homeomorphism ϕ_N of \mathbb{R}^k , which is close to the identity map, the function $g_N(x) = g(\phi_N(x))|\det D\phi_N(x)|$ remains close in L^1 to g . Lemma 12 relaxes the Lipschitz assumption and allows us to take functions in $\text{ALip}(r)$.

Lemma 16. *Let $g \in L^1(\mathbb{R}^k)$ be Lipschitz and satisfy the following tail bounds for some $\mathcal{M} > 0, \beta > 0$:*

$$\int_{B(0,R)^c} |g(x)|dx \leq \mathcal{M}R^{-\beta}.$$

Let $\{\phi_N : \mathbb{R}^k \rightarrow \mathbb{R}^k\}$ be a sequence of continuous, 1-1 and onto functions, have a (matrix-valued) derivatives $D\phi_N$ defined a.e., and

$$|\phi_N(x) - x| \leq \epsilon_N|x| \text{ a.e. } x, \|D\phi_N - \text{Id}\| \leq \epsilon_N \tag{3.16}$$

for some ϵ_N which approaches 0. Define g_N via

$$g_N(x) = g(\phi_N(x))|\det D\phi_N(x)|.$$

Then we can quantitatively show that $\lim_{N \rightarrow \infty} \|g - g_N\|_{L^1} = 0$:

$$\begin{aligned} \|g - g_N\|_{L^1(\mathbb{R}^k)} \leq & (k + 1 + \beta) \left(\frac{2\mathcal{M}}{(k+1)(1-\epsilon_N)^{k+\beta}} \right)^{\frac{k+1}{k+1+\beta}} \left(\frac{\|g\|_{\text{Lip}} |S^{k-1}(1)|}{\beta} \right)^{\frac{\beta}{k+1+\beta}} \epsilon_N^{\frac{\beta}{k+1+\beta}} \\ & + \frac{(1+\epsilon_N)^k - 1}{(1-\epsilon_N)^k} \int_{\mathbb{R}^k} |g(y)| dy \end{aligned} \tag{3.17}$$

In (3.16) $|\cdot|$ denotes the Euclidean norm on \mathbb{R}^k and $\|\cdot\|$ the associated Matrix norm, given by $\|M\| = \sup\{|Mx| : |x| = 1\}$.

Proof. In evaluating $\|g - g_N\|_{L^1(\mathbb{R}^k)}$, we add and subtract the term $g(\phi_N(x))$ to show that $\|g - g_N\|_{L^1(\mathbb{R}^k)}$ is bounded by a sum of two terms:

$$\int_{\mathbb{R}^k} |g(x) - g(\phi_N(x))|dx + \int_{\mathbb{R}^k} |g(\phi_N(x))| |1 - \det(D\phi_N(x))| dx. \tag{3.18}$$

We control the first term in (3.18) as follows. When x is small ($|x| \leq R$) we use that fact that g is Lipschitz and that the first part of (3.16) holds, to get the upper bound $\|g\|_{\text{Lip}}\epsilon_N \int_{B(0,R)} |x|dx$.

When $|x| > R$, we let $T_R = \int_{B(0,R)^c} |g(x)|dx$. We bound $|g(x) - g(\phi_N(x))|$ by $|g(x)| + |g(\phi_N(x))|$. The integral of this last term is bounded above by

$$T_R + \left| \sup_{x' \in \mathbb{R}^k} \frac{1}{\det(D\phi_N(x'))} \right| \int_{B(0,R)^c} |g(\phi_N(x))| \det(D\phi_N(x)) dx. \tag{3.19}$$

The second half of (3.16) implies that $\det(D\phi_N(x)) \in [(1 - \epsilon_N)^k, (1 + \epsilon_N)^k]$. This, together with the observation that $|x| > R \Rightarrow |\phi_N(x)| \geq (1 - \epsilon_N)R$ implies that the expression in (3.19) is at most

$$T_R + \frac{1}{(1 - \epsilon_N)^k} T_{(1-\epsilon_N)R}.$$

We now treat the second term in (3.18). It is bounded above by

$$\sup_{x' \in \mathbb{R}^k} \frac{|1 - \det(D\phi_N(x'))|}{\det(D\phi_N(x'))} \int_{\mathbb{R}^k} |g(\phi_N(x))| \det(D\phi_N(x)) dx,$$

which, by (3.16) and the observation that $(1 + \epsilon_N)^k - 1 \geq 1 - (1 - \epsilon_N)^k$, can be seen to be bounded above by

$$\frac{(1 + \epsilon_N)^k - 1}{(1 - \epsilon_N)^k} \int_{\mathbb{R}^k} |g(y)| dy.$$

In summary, we have shown that

$$\begin{aligned} \|g - g_N\|_{L^1(\mathbb{R}^k)} &\leq \|g\|_{\text{Lip}} \frac{|S^{k-1}(1)|}{k+1} \epsilon_N R^{k+1} + \frac{2}{(1 - \epsilon_N)^k} T_{(1-\epsilon_N)R} \\ &\quad + \frac{(1 + \epsilon_N)^k - 1}{(1 - \epsilon_N)^k} \int_{\mathbb{R}^k} |g(y)| dy. \end{aligned} \tag{3.20}$$

Equation (3.17) follows from (3.20) by using $T_{(1-\epsilon_N)R} = \mathcal{M}(1 - \epsilon_N)^\beta R^\beta$ and choosing

$$R^{k+1+\beta} = \frac{2\mathcal{M}\beta}{\|g\|_{\text{Lip}} |S^{k-1}(1)| (1 - \epsilon_N)^{k+\beta} \epsilon_N}. \quad \square$$

Lemma 17. *Let k, ϕ_N, ϵ_N be as in Lemma 16. Let $f \in \text{ALip}(r, L)$ for some (r, L) satisfy*

$$\int_{B(0,R)^c} |f(x)| dx \leq \mathcal{M}R^{-\beta},$$

and let $f_N(x) = f(\phi_N(x)) |\det D\phi_N(x)|$. Then

$$\|f - f_N\|_{L^1(\mathbb{R}^k)} \leq C(k, \beta, \mathcal{M}, L) \epsilon_N^{1/(1+r+(k+1)/\beta)}.$$

The constant can be made explicit.

Proof. Given ϵ , we can choose g as in the hypothesis, which satisfies the same tail bound T_R up to a factor of 2. By a 3ϵ argument, we have

$$\|f - f_N\|_{L^1(\mathbb{R}^k)} \leq 2\epsilon + \begin{pmatrix} \text{r.h.s. of (3.17)} \\ \text{with } 2\mathcal{M} \\ \text{and } \|g\|_{\text{Lip}} = L\epsilon^{-r} \end{pmatrix}$$

and we choose $\epsilon \sim \epsilon_N^{1/(1+r+(k+1)/\beta)}$. □

Lemma 18. *Fix $k \in \mathbb{N}$ and $\delta \in (0, 2]$. Let $\eta_1(q), \eta_2(q)$, and $\eta_3(q)$ be given by*

$$\eta_1(q) = \frac{2 + \delta}{k + 3 + \delta} q, \quad \eta_2(q) = \frac{\delta}{2} - q \left(1 + \frac{\delta}{2}\right), \quad \text{and } \eta_3(q) = \frac{2 + \delta}{4} (1 - q).$$

Then

$$\max_{q \geq 0} [\min(\eta_1(q), \eta_2(q), \eta_3(q))] = \frac{\delta}{k + 5 + \delta + r(2 + \delta)} \tag{3.21}$$

which is $\eta_1(q_*)$ with

$$q_* = \left(\frac{\delta}{\delta + 2}\right) \left(\frac{k + 3 + \delta + r(2 + \delta)}{k + 5 + \delta + r(2 + \delta)}\right) \tag{3.22}$$

Proof. Since $\delta \leq 2$, it is easy to show $\eta_2(q) \leq \eta_3(q)$ on $[0, 1]$. Thus, $\eta_3(q)$ can be neglected in the left hand side of (3.21). We are left with maximizing the minimum of two lines: one increasing and the other decreasing. Thus, this maximum is at their intersection. This provides q_* in (3.22) and the maximum in (3.21). □

We now show how non-smooth can the functions in $\text{ALip}(r)$ be, by computing the exponent r in the cost in the Lipschitz in the definition 9 for some functions in $L^1_{\text{loc}}(\mathbb{R})$.

Lemma 19. *The following table gives the cost in terms of the Lipschitz constant, of ϵ -approximating some functions using the L^1 metric.*

function	$\ g_\epsilon\ _{Lip}$ (upper bound)
$ x ^a, a \in (-1, 1]$	$C(a)\epsilon^{-\frac{1-a}{1+a}}$
$\mathbf{1}_{[0,\infty)}$	$\frac{1}{4}\epsilon^{-1}$
$f \in C^\alpha \cap L^1((1 + x ^\beta)dx)$	$C(f)\epsilon^{-\frac{1+\beta}{\alpha\beta}}$

Proof. We show the proofs of some of them.

1. When $f(x) = |x|^a$, with $a \in (-1, 0)$. Set

$$f_h(x) = \begin{cases} |x|^a, & |x| \geq h \\ h^a + ah^{a-1}(x-h), & 0 \leq |x| \leq h \end{cases}$$

Then $\|f - f_h\|_{L^1} = 2a \left[\frac{1}{1-a} + \frac{1}{2} \right] h^{1-a} =: \epsilon$, while $\|f_h\|_{Lip} = 2h^{-(1+a)} = C(a)\epsilon^{-\frac{1+a}{1-a}}$. If $a > 0$, the formula for f_h above need not change.

2. When $f(x) = \mathbf{1}_{[0,\infty)}$, we set $f_h(x) = \begin{cases} 1, & x > h/2, \\ \frac{h}{2} + \frac{2-h}{h}x, & -h/2 < x < h/2 \\ 0, & \text{otherwise} \end{cases}$

Then $\|f - f_h\|_{L^1} = (h/4) =: \epsilon$, while $\|f_h\|_{Lip} = (1/h)$.

3. Let f be Hölder continuous with order α , and have tail bounds

$$\int_{[-R,R]^c} |f| dx \leq MR^{-\beta}.$$

It easily follows that $\|f\|_{L^\infty} < \infty$, and an elementary computation shows that, whenever $|h| \leq 1$, we have

$$\|\tau_h f - f\|_{L^1} \leq 2 \max \left\{ \|f\|_\alpha, \|f\|_\alpha^{\frac{\beta}{1+\beta}} \right\} \left[\frac{1}{\beta} + (\beta M)^{\frac{1}{1+\beta}} \right] |h|^{\frac{\beta}{1+\beta}\alpha}.$$

Here $\tau_h f(x)$ is the translate of f ($= f(x-h)$). Using this inequality, choose $\psi \geq 0 \in C^1(\mathbb{R})$ that satisfies $\int_{\mathbb{R}} \psi = 1$ and $\text{supp}(\psi) \subset [-1, 1]$. Let $\psi_\delta(x) = \delta^{-1}\psi(x/\delta)$. Then, for any $\delta > 0$, $f * \psi_\delta$ is Lipschitz and we have:

$$\begin{aligned} \|f * \psi_\delta - f\|_{L^1} &\leq \sup_{|h| \leq \delta} \|\tau_h f - f\|_1 = C(\|f\|_{C^\alpha}, \mathcal{M}, \beta) \delta^{\frac{\beta\alpha}{1+\beta}}, \\ \|(f * \psi_\delta)'\|_{L^\infty} &\leq \int_{\mathbb{R}} |f(x-y)\delta^{-2}\psi'(y/\delta)| dy \leq \delta^{-1}\|\psi'\|_{L^1}\|f\|_{L^\infty}. \end{aligned}$$

Hence, choosing δ so that $\|f * \psi_\delta - f\|_{L^1} = \epsilon$, gives $\|(f * \psi_\delta)'\|_{L^\infty} \leq C\epsilon^{-(1+\beta)/(\alpha\beta)}$. \square

Lemma 20. *For any $r > 0$ and $\beta > 0$ the following inclusion holds for $\text{ALip}(r)$:*

$$\text{ALip}(r) \cap \left\{ f : \int_{\mathbb{R}} |x|^\beta |f(x)| dx < \infty \right\} \subseteq B_{1,\infty}^q(\mathbb{R}),$$

where $q = \frac{\beta}{1+\beta(1+r)}$ and $B_{1,\infty}^q(\mathbb{R})$ is a Besov space².

²We recall that $B_{a,\infty}^q$ can be characterized as the space of functions $f \in L^a$ such that

$$\sup_{h>0} h^{-q} \|f - \tau_h f\|_{L^a} < \infty.$$

The proof is omitted, since it is standard and involves computations similar to those in the proof of 19.

We close this section by giving a remark concerning the strength of our POC result in L^1 , whenever it applies, compared to the weak- L^1 propagation of chaos result for $\{\hat{f}^N\}_N$ which can be obtained via the Dunford-Pettis theorem whenever $H[f|\gamma] < \infty$, as hinted at in [3].

Remark 21. If $H[f|\gamma] < \infty$ holds, it will follow from Theorem 27 below that

$$\sup_N N^{-1} H[\hat{f}^N | \sigma^N] < \infty.$$

As remarked in [3], this implies that $\sup_N H(\Pi_k \hat{f}^N | \gamma^{\otimes k}) < \infty$ (For the proof, see [2] when $k = 1$ and see [1] for general k). Thus, the sequence $\{\Pi_k \hat{f}^N\}_{N \geq k+1}$ is uniformly integrable and the Dunford-Pettis theorem implies that $\{\Pi_k \hat{f}^N\}_{N \geq k+1}$ is compact in the weak topology. Thus, $\Pi_k \hat{f}^N$ weakly converges to $f^{\otimes k}$ in L^1 (i.e. against all $\phi \in L^\infty(\mathbb{R}^k)$ and not just for ϕ continuous and bounded). Theorem 10, which holds for $f \in \text{ALip}(r)$ has two advantages over this result: it is quantitative and gives convergence in the strong L^1 norm.

4 Entropic chaos

We now study entropic chaos for rescaled tensor products. Let us first recall Theorem 12 in [3] (see also [5, Theorem 1.15]). For convenience, we provide a proof in the Appendix. Our proof will avoid the conditioned states and the associated local version of the central limit theorem used in [3], and rely instead on rescaling and the classical central limit theorem.

Theorem 22 (asymptotic upper semi-extensivity of the entropy). *For each $N \in \mathbb{N}$, let $F^N \in \mathcal{P}_{\text{sym}}(\mathcal{K}^N)$ be such that $\lim_N \Pi_1 F^N = f$ weakly, for some $f \in \mathcal{P}(\mathbb{R})$ satisfying $H(f|\gamma) < \infty$. Then,*

$$H(f|\gamma) \leq \liminf_{N \rightarrow \infty} \frac{1}{N} H(F^N | \sigma^N). \tag{4.1}$$

In view of (4.1), we see that, in order to prove that an f -chaotic sequence F^N is also f -entropically chaotic, it suffices to show that $\limsup_N N^{-1} H(F^N | \sigma^N) \leq H(f|\gamma)$. We will use this strategy to prove that \hat{f}^N is entropically chaotic to f , which is the main goal of this section. We will need the following formula for the density of the rescaled measure:

Proposition 23 (formula for rescaled densities). *Let $F^N \in \mathcal{P}(\mathbb{R}^N)$ have a density. Then $\hat{F}^N \ll \sigma^N$, and we have:*

$$\begin{aligned} \frac{d\hat{F}^N}{d\sigma^N}(x) &= \sqrt{N} |\mathcal{K}^N| \int_0^\infty r^{N-1} F^N(rx) dr, \\ &= |S^{N-1}(1)| \int_0^\infty r^{N-1} F^N(rx/|x|) dr \quad \forall x \in \mathcal{K}^N. \end{aligned} \tag{4.2}$$

Proof. For any test function $\phi : \mathcal{K}^N \rightarrow \mathbb{R}$, using polar coordinates, we have:

$$\begin{aligned} \int_{\mathcal{K}^N} \phi(x) \hat{F}^N(dx) &= \int_{\mathbb{R}^N} \phi(\hat{x}) F^N(x) dx \\ &= \int_{S^{N-1}(1)} \int_0^\infty r^{N-1} \phi(\sqrt{N}\omega) F^N(r\omega) dr d\omega \\ &= (\sqrt{N})^N \int_{S^{N-1}(1)} \phi(\sqrt{N}\omega) \int_0^\infty u^{N-1} F^N(u\sqrt{N}\omega) du d\omega \end{aligned}$$

$$= \sqrt{N} \int_{\mathcal{K}^N} \phi(y) \int_0^\infty u^{N-1} F^N(uy) du dy,$$

where we have used the changes of variables $r = \sqrt{N}u$ and $y = \sqrt{N}\omega$. The conclusion now follows simply noting that $dy = |\mathcal{K}^N| \sigma^N(dy)$. The formula in (4.2) follows from a simple rescaling. \square

In Lemma 26 below, we write the difference between $H(\hat{F}^N | \sigma^N)$ and $H(F^N | \gamma^{\otimes N})$ as the negative of some relative entropy with respect to the following distribution:

Definition 24 (angular version). *Let $F^N \in \mathcal{P}(\mathbb{R}^N)$ without an atom at 0. We define an angular version of F^N , denoted $\check{F}^N \in \mathcal{P}(\mathbb{R}^N)$, as the law of $|Z^N| \hat{X}^N / \sqrt{N}$, where $X^N \sim F^N$ and $Z^N \sim \gamma^{\otimes N}$ are independent.*

The following lemma provides a formula for the density of the angular version in terms of the density of the rescaled measure:

Lemma 25 (formula for the density of the angular version). *Let $F^N \in \mathcal{P}(\mathbb{R}^N)$ be such that \hat{F}^N has a density with respect to σ^N . Then \check{F}^N has a density with respect to the Lebesgue measure, given by*

$$\check{F}^N(x) = \gamma^{\otimes N}(x) \frac{d\hat{F}^N}{d\sigma^N}(\hat{x}), \quad \forall x \in \mathbb{R}^N. \tag{4.3}$$

Proof. For any test function $\phi : \mathbb{R}^N \rightarrow \mathbb{R}$, we have

$$\begin{aligned} \int_{\mathbb{R}^N} \phi(x) \check{F}^N(dx) &= \int_{\mathbb{R}^N} \int_{\mathcal{K}^N} \phi\left(\frac{|z|y}{\sqrt{N}}\right) \hat{F}^N(dy) \gamma^{\otimes N}(z) dz \\ &= \int_{S^{N-1}(1)} \int_0^\infty \int_{\mathcal{K}^N} \phi\left(\frac{ry}{\sqrt{N}}\right) \hat{F}^N(dy) \gamma^{\otimes N}(r\omega) r^{N-1} dr d\omega, \end{aligned}$$

where we have changed from Cartesian $z \in \mathbb{R}^N$ to polar coordinates $(r, \omega) \in (0, \infty) \times S^{N-1}(1)$. The key step is to note that $\gamma^{\otimes N}(r\omega) = \gamma^{\otimes N}(ry/\sqrt{N})$ for any $y \in \mathcal{K}^N$ and $\omega \in S^{N-1}(1)$. With this, and noting that then the integrand does not depend on ω , we obtain:

$$\int_{\mathbb{R}^N} \phi(x) \check{F}^N(dx) = |S^{N-1}(1)| \int_0^\infty \int_{\mathcal{K}^N} \phi\left(\frac{ry}{\sqrt{N}}\right) \hat{F}^N(dy) \gamma^{\otimes N}\left(\frac{ry}{\sqrt{N}}\right) r^{N-1} dr.$$

Since \hat{F}^N has a density with respect to σ^N , we have

$$\hat{F}^N(dy) = \frac{d\hat{F}^N}{d\sigma^N}(y) \sigma^N(dy) = \frac{d\hat{F}^N}{d\sigma^N}(\sqrt{N}\omega) \frac{d\omega}{|S^{N-1}(1)|},$$

for $\omega = y/\sqrt{N} \in S^{N-1}(1)$. Consequently,

$$\begin{aligned} \int_{\mathbb{R}^N} \phi(x) \check{F}^N(dx) &= \int_0^\infty \int_{S^{N-1}(1)} \phi(r\omega) \frac{d\hat{F}^N}{d\sigma^N}(\sqrt{N}\omega) d\omega \gamma^{\otimes N}(r\omega) r^{N-1} dr \\ &= \int_{\mathbb{R}^N} \phi(x) \frac{d\hat{F}^N}{d\sigma^N}(\hat{x}) \gamma^{\otimes N}(x) dx, \end{aligned}$$

where we changed back to Cartesian coordinates $x = r\omega \in \mathbb{R}^N$. The result follows. \square

Lemma 26. *Let F^N be a probability density on \mathbb{R}^N such that $H(F^N | \gamma^{\otimes N}) < \infty$. Then,*

$$H(\hat{F}^N | \sigma^N) - H(F^N | \gamma^{\otimes N}) = -H(F^N | \check{F}^N) \leq 0. \tag{4.4}$$

Proof. We have:

$$\begin{aligned} H(\hat{F}^N | \sigma^N) - H(F^N | \gamma^{\otimes N}) &= \int_{\mathbb{R}^N} \log \left(\frac{d\hat{F}^N}{d\sigma^N}(\hat{x}) \right) F^N(x) dx - \int_{\mathbb{R}^N} \log \left(\frac{F^N(x)}{\gamma^{\otimes N}(x)} \right) F^N(x) dx \\ &= \int_{\mathbb{R}^N} \log \left(\frac{\check{F}^N(x)}{F^N(x)} \right) F^N(x) dx, \end{aligned}$$

where we have used (4.3). The last expression equals $-H(F^N | \check{F}^N)$. \square

We can now state and prove one of our main results: entropic chaoticity for rescaled tensor products, under the minimal assumptions of unit energy and finite entropy relative to γ .

Theorem 27 (entropic chaos for rescaled tensor products). *Let f be a probability density on \mathbb{R} such that $\int_{\mathbb{R}} x^2 f(x) dx = 1$ and $H(f | \gamma) < \infty$. Then \hat{f}^N is entropically chaotic to f .*

Proof. Since $H(f^{\otimes N} | \gamma^{\otimes N}) = NH(f | \gamma)$, taking $F^N = f^{\otimes N}$ in (4.4) gives $\frac{1}{N}H(\hat{f}^N | \sigma^N) \leq H(f | \gamma)$. Taking \limsup_N and using (4.1), the result follows (recall that, since the second moment of f is equal to 1, we know that \hat{f}^N is f -chaotic, thanks to Theorem 7). \square

The last result of this section provides a quantitative entropic chaos rate for \hat{f}^N , see Theorem 32 below. The proof relies on Proposition 36 provided in the next section, and on the following result, which is a slight improvement of [10, Theorem 4.17]:

Theorem 28. *Let $F^N \in \mathcal{P}_{\text{sym}}(\mathcal{K}^N)$ be an f -chaotic sequence, for some $f \in \mathcal{P}(\mathbb{R}) \cap L^p(\mathbb{R})$ for some $p > 1$. Assume that there exists $M > 0$ such $\int_{\mathbb{R}^N} |x_1|^k F^N(dx) \leq M$ for some $k > 4$, and that $\frac{1}{N}I(F^N | \sigma^N) \leq M$, for all N . Then F^N is entropically chaotic to f , with the following explicit rate:*

$$\left| \frac{1}{N}H(F^N | \sigma^N) - H(f | \gamma) \right| \leq C_1 \left(N^{-1/2}W_2(F^N, f^{\otimes N}) + C_2N^{-\eta} + C_3N^{-(k/4-1)} \right), \quad (4.5)$$

for any $\eta < \frac{1}{8} \frac{k-2}{k+1}$, where $C_1 = C_1(M)$, $C_2 = C_2(\eta)$, and $C_3 = C_3(p, \|f\|_p, k, M)$.

The authors in [10] require at least 6 finite moments on f and, although they present their theorem in terms of a normalized W_1 metric, they provide the above result for W_2 in their proof. We manage to relax this condition to $k > 4$ finite moments. To achieve this, we analyze the method in their proof. We will need to define the conditioned tensor product states.

Definition 29. *Let $f \in L^1(\mathbb{R})$. Then the conditioned state $[f^{\otimes N}]_N$ is the element in $L^1(\mathcal{K}^N, \sigma^N)$ given by*

$$[f^{\otimes N}]_N(v) = \frac{f^{\otimes N}(v)}{\int_{\mathcal{K}^N} f^{\otimes N}(y) \sigma^N(dy)}.$$

We note that when $f \in L^1(\mathbb{R})$, the product structure of $f^{\otimes N}$ makes $[f^{\otimes N}]_N$ well defined, in spite of the denominator in its definition depending on the values of $f^{\otimes N}$ on a set of measure zero (see the note in the beginning of the proof of Theorem 4.9 in [10]).

The proof of Theorem [10, Theorem 4.17] relies on two results concerning the conditioned tensor product states. The first one concerns their entropic chaoticity: when $f \in \mathcal{P}(\mathbb{R}) \cap L^p(\mathbb{R}) \cap \mathcal{P}_6(\mathbb{R})$, we have

$$\left| \frac{1}{N}H([f^{\otimes N}]_N | \sigma^N) - H(f | \gamma) \right| \leq C(p, \|f\|_p, M)N^{-1/2}. \quad (4.6)$$

The second result gives a uniform bound on their Fisher information: if $f \in \mathcal{P}(\mathbb{R}) \cap L^p(\mathbb{R}) \cap \mathcal{P}_6(\mathbb{R})$ and $I(f|\gamma) < \infty$, then $\sup_N \frac{1}{N} I([f^{\otimes N}]_N | \sigma^N)$ (see Theorem 4.14 in [10]). In the following two lemmas we extend these results to the case $f \in \mathcal{P}_k(\mathbb{R})$ with $k = 4 + r$ for some $r \in (0, 2]$.

Lemma 30. *Let $f \in \mathcal{P}_{4+r}(\mathbb{R}) \cap L^p(\mathbb{R})$ for some $r \in (0, 2], p > 1, \int_{\mathbb{R}} v^2 f(v) dv = 1$. Then*

$$\left| \frac{1}{N} H([f^{\otimes N}]_N | \sigma^N) - H(f|\gamma) \right| \leq C \left(p, \|f\|_p, \int_{\mathbb{R}} |v|^{4+r} f(v) dv \right) N^{-r/4}. \quad (4.7)$$

We note that we do not require $\int_{\mathbb{R}} v f(v) dv$ to be 0. The proof of Lemma 30 requires only small adjustments to that of (4.6), which we list in the Appendix.

Lemma 31. *Let $f \in \mathcal{P}(\mathbb{R}) \cap L^p(\mathbb{R})$ for some $p > 1, \int_{\mathbb{R}} |v|^2 f(v) dv = 1$, and assume that $I(f|\gamma) < \infty$. Then*

$$\sup_{N \geq 2} \frac{1}{N} I([f^{\otimes N}]_N | \sigma^N) < \infty.$$

Lemma 31 follows from the proof of the first statement in [10, Theorem 4.14], and noting that we do not need the higher moment assumptions on f for this part. For completeness, we sketch the proof in the Appendix. We are now ready to prove Theorem 28, which is an adaptation to that of [10, Theorem 4.17].

Proof of Theorem 28. Let F^N and f be as in the hypotheses. A version of the HWI inequality on spaces of positive Ricci curvature, proven in [14] (see also [16, Theorem 30.21]), implies that

$$\frac{1}{N} |H(F^N | \sigma^N) - H([f^{\otimes N}]_N | \sigma^N)| \leq \frac{\pi}{2} \sqrt{\frac{\max\{I([f^{\otimes N}]_N, \sigma^N), I(F^N | \sigma^N)\}}{N} \frac{W_2(F^N, [f^{\otimes N}]_N)}{\sqrt{N}}}.$$

Thus, we have that

$$\begin{aligned} \left| \frac{1}{N} H(F^N | \sigma^N) - H(f|\gamma) \right| &\leq \frac{1}{N} |H(F^N | \sigma^N) - H([f^{\otimes N}]_N | \sigma^N)| + \\ &\quad \left| \frac{1}{N} H([f^{\otimes N}]_N | \sigma^N) - H(f|\gamma) \right| \\ &\leq C_1 \frac{W_2(F^N, [f^{\otimes N}]_N)}{\sqrt{N}} + \left| \frac{1}{N} H([f^{\otimes N}]_N | \sigma^N) - H(f|\gamma) \right| \\ &\leq C_1 \left(\frac{W_2(F^N, f^{\otimes N}) + W_2(f^{\otimes N}, [f^{\otimes N}]_N)}{\sqrt{N}} \right) \\ &\quad + \left| \frac{1}{N} H([f^{\otimes N}]_N | \sigma^N) - H(f|\gamma) \right| \end{aligned}$$

where we used the triangle inequality in the first and last steps, and we used Proposition 36 (below) and Lemma 31 to bound $\frac{1}{N} I(\hat{f}^N | \sigma^N)$ and $\frac{1}{N} I([f^{\otimes N}]_N | \sigma^N)$. This, together with Lemma 30 and the bound

$$\frac{1}{\sqrt{N}} W_2([f^{\otimes N}]_N, f^{\otimes N}) \leq C(\eta) N^{-\eta}$$

for any $\eta < \frac{1}{8} \frac{k-2}{k+1}$, provided in the proof of [10, Theorem 4.17], proves (4.5). \square

We are now ready to state and prove the main theorem of this section.

Theorem 32 (quantitative entropic chaos for rescaled tensor products). *Let $f \in \mathcal{P}_k(\mathbb{R}) \cap L^p(\mathbb{R})$ for some $k > 4, p > 1$. Assume that $\int_{\mathbb{R}} x^2 f(x) dx = 1$ and $I(f|\gamma) < \infty$. Then, \hat{f}^N is entropically chaotic to f , with the following explicit rate: for any $\eta < \frac{1}{8} \frac{k-2}{k+1}$, there exists a constant $C > 0$ such that*

$$\left| \frac{1}{N} H(\hat{f}^N | \sigma^N) - H(f|\gamma) \right| \leq C \left(N^{-\eta} + N^{-(k/4-1)} \right).$$

Proof. By Theorem 7, \hat{f}^N is f -chaotic; by Lemma 6, the k -th moment of \hat{f}^N is bounded; and by Proposition 36 we know that $\frac{1}{N} I(\hat{f}^N | \sigma^N)$ is bounded. Thus, Theorem 28 gives (4.5). Thanks to Theorem 7, $N^{-1/2} W_2(\hat{f}^N, f^{\otimes N}) \leq CN^{-1/4}$. Since $\frac{1}{4} > \frac{1}{8} \frac{k-2}{k+1}$ for all $k > 0$, the $N^{-1/4}$ term can be neglected. The result follows. \square

5 Fisher information chaos

In this section we prove *Fisher information chaos*, which we now define, for the rescaled states $\{\hat{f}^N\}$.

Definition 33 (Fisher information chaos). *For each $N \in \mathbb{N}$, let $F^N \in \mathcal{P}_{\text{sym}}(\mathcal{K}^N)$, and let $f \in \mathcal{P}(\mathbb{R})$ such that $I(f|\gamma) < \infty$. The sequence $(F^N)_{N \in \mathbb{N}}$ is said to be Fisher information chaotic to f if it is Kac chaotic to f , and*

$$\lim_{N \rightarrow \infty} \frac{1}{N} I(F^N | \sigma^N) = I(f|\gamma). \tag{5.1}$$

The main result of this section is the following:

Theorem 34 (Fisher information chaos for rescaled tensor products). *Let f be a probability density on \mathbb{R} such that $\int_{\mathbb{R}} x^2 f(x) dx = 1$ and $I(f|\gamma) < \infty$. Then \hat{f}^N is Fisher information chaotic to f .*

To prove this theorem, we need to establish the equality in (5.1). The new part is in obtaining the limit inequality

$$\limsup_{N \rightarrow \infty} N^{-1} I(\hat{f}^N | \sigma^N) \leq I(f|\gamma),$$

which we will treat in Proposition 36. The proof uses the convexity of the map $(x, \vec{y}) \mapsto \frac{|\vec{y}|^2}{x}$. We prove a slightly more general result in Proposition 35. This proposition depends on the technical, but not too difficult, Lemmas 37-39. The reverse inequality, i.e., $\liminf_{N \rightarrow \infty} N^{-1} I(\hat{f}^N | \sigma^N) \geq I(f|\gamma)$, has been proven in [10, Theorem 4.15] for sequences more general than \hat{f}^N .

Proposition 35. *Let $F \in L^1(\mathbb{R}^N, d^N x)$ and $G \in L^1(\mathcal{K}^N, \sigma^N)$ be probability densities; here $N \geq 3$. Then the following hold.*

1. $I(G|\sigma^N) = \frac{N-2}{N} I(\check{G}|\gamma^{\otimes N})$.
2. $I(\check{F}|\gamma^{\otimes N}) \leq \frac{1}{(N-2)} \int_{\mathbb{R}^N} \gamma^{\otimes N}(x) \frac{|x|^2 |\nabla \rho|^2(x) - (\nabla \rho(x) \cdot x)^2}{\rho(x)} dx$, where $\rho(x) = F(x)/\gamma^{\otimes N}(x)$

Proposition 36. *Let f be a probability density on \mathbb{R} such that $\int_{\mathbb{R}} x^2 f(x) dx = 1$ and $I(f|\gamma) < \infty$. Then for all $N \geq 2$ we have the inequality:*

$$\frac{1}{N} I(\hat{f}^N | \sigma^N) \leq \left(1 - \frac{1}{N} \right) I(f|\gamma)$$

Proof of Theorem 34. Proposition 36 implies that $\limsup_{N \rightarrow \infty} N^{-1} I(\hat{f}^N | \sigma^N) \leq I(f|\gamma)$. As mentioned above, the reverse inequality: $\liminf_{N \rightarrow \infty} N^{-1} I(\hat{f}^N | \sigma^N) \geq I(f|\gamma)$ has been proven in [10, Theorem 4.15]. This establishes the Fisher information chaoticity for the family $\{\hat{f}^N\}_N$. \square

Before proving Proposition 35, we provide two useful technical lemmas. The first one gives a formula for the gradient of a function $\tilde{J} : \mathbb{R}^N \setminus \{0\} \rightarrow \mathbb{R}$ that depends only on $x/|x|$, while the second lemma reduces the integral of such a function integrated against a Gaussian, to an integral on the unit sphere. We omit the proof of the first lemma since it follows directly from the chain rule.

Lemma 37 (gradient of the composition). *Let $J : \mathbb{R}^N \rightarrow \mathbb{R}$ be C^1 , let $r > 0$ and $x \neq 0$. Let \tilde{J} be the function on $\mathbb{R}^N \setminus \{0\}$ defined by*

$$\tilde{J}(x) = J\left(r \frac{x}{|x|}\right).$$

Then, for $x \neq 0$,

$$\nabla \tilde{J}(x) = \frac{r}{|x|} \left(\text{Id} - \frac{xx^T}{|x|^2} \right) [\nabla J] \left(r \frac{x}{|x|} \right)$$

In particular, because the matrix $\text{Id} - \frac{xx^T}{|x|^2}$ is a projection, we have

$$|\nabla \tilde{J}(x)|^2 = \frac{r^2}{|x|^2} \left(\left| [\nabla J] \left(r \frac{x}{|x|} \right) \right|^2 - \left([\nabla J] \left(r \frac{x}{|x|} \right) \cdot \frac{x}{|x|} \right)^2 \right)$$

Lemma 38. *Let $N \geq 3$ and let $\phi \in C(\mathbb{R}^N \setminus \{0\})$ be any function that depends only on $x/|x|$, then*

$$\int_{\mathbb{R}^N} |x|^{-2} \gamma^{\otimes N}(x) \phi\left(\frac{x}{|x|}\right) dx = \frac{1}{(N-2)} \int_{S^{N-1}(1)} \phi(w) \sigma_1^N(dw).$$

Proof. Using the polar coordinates ($x = rw, dx \leftrightarrow r^{n-1} dr |S^{N-1}(1)| \sigma_1^N(dw)$) we see that

$$\int_{\mathbb{R}^N} |x|^{-2} \gamma^{\otimes N}(x) \phi\left(\frac{x}{|x|}\right) dx = \int_{S^{N-1}(1)} \phi(w) \sigma_1^N(dw) |S^{N-1}(1)| \int_{r=0}^{\infty} r^{N-3} \frac{e^{-\frac{r^2}{2}}}{(2\pi)^{\frac{N}{2}}} dr.$$

Recall that $\int_0^\infty r^{N-3} e^{-\frac{r^2}{2}} dr = \frac{(2\pi)^{\frac{N}{2}-1}}{|S^{N-3}(1)|}$, and that $|S^{N-1}(1)| = \frac{2\pi^{\frac{N}{2}}}{\Gamma(\frac{N}{2})}$. Hence

$$\begin{aligned} |S^{N-1}(1)| \int_{r=0}^{\infty} r^{N-3} \frac{e^{-\frac{r^2}{2}}}{(2\pi)^{\frac{N}{2}}} dr &= \frac{1}{2\pi} \frac{|S^{N-1}(1)|}{|S^{N-3}(1)|} \\ &= \frac{1}{2\pi} \frac{\pi}{(N/2-1)} = \frac{1}{N-2}. \quad \square \end{aligned}$$

We are now ready to prove Proposition 35.

Proof of Proposition 35. Note that $\check{G}(x)/\gamma^{\otimes N}(x) = G(\sqrt{N} \frac{x}{|x|})$.

Notice that $\nabla G(\sqrt{N} \frac{x}{|x|}) = \frac{\sqrt{N}}{|x|} \nabla_S G|_{\sqrt{N} \frac{x}{|x|}}$. Thus, we have:

$$\begin{aligned} I(\check{G}|\gamma^{\otimes N}) &= \int_{\mathbb{R}^N} \gamma^{\otimes N}(x) \frac{N}{|x|^2} \frac{\left(\nabla_S G \left(\sqrt{N} \frac{x}{|x|} \right) \right)^2}{G \left(\sqrt{N} \frac{x}{|x|} \right)} d^N x \\ &= \frac{N}{N-2} \int_{\mathcal{K}^N} \sigma^N(dw) \frac{|\nabla_S G(w)|^2}{H(w)} \\ &= \frac{N}{N-2} I(G|\sigma^N). \end{aligned}$$

Here we used Lemmas 38 and 37. This proves point 1.

In order to prove point 2, we will resort to Jensen's inequality. Recall that the mapping

$$(x, \vec{y}) \mapsto \frac{|\vec{y}|^2}{x}$$

from $\mathbb{R} \setminus \{0\} \times \mathbb{R}^d \rightarrow \mathbb{R}$ (for any d) is convex, also the following representation formula for $\hat{F}(x)$ (here $|x| = R$)

$$\hat{F}(x) = \int_{r=0}^{\infty} r^{N-1} \frac{F}{\gamma^{\otimes N}} \left(r \frac{x}{|x|} \right) \gamma^N(r) |S^{N-1}(1)| dr,$$

which is obtained from (4.2), convex since $\int_{r=0}^{\infty} r^{N-1} \gamma^N(r) |S^{N-1}(1)| dr = 1$. The Fisher information of \check{F} is given by

$$\begin{aligned} I(\check{F}|\gamma^{\otimes N}) &= \int_{\mathbb{R}^N} \gamma^{\otimes N}(x) \frac{|S^{N-1}(1)|^2 |\nabla \int_0^\infty r^{N-1} F(r \frac{x}{|x|}) dr|^2}{|S^{N-1}(1)| \int_0^\infty r^{N-1} F(r \frac{x}{|x|}) dr} dx \\ &= \int_{\mathbb{R}^N} \gamma^{\otimes N}(x) \frac{|\nabla \int_0^\infty r^{N-1} \frac{F}{\gamma^{\otimes N}}(r \frac{x}{|x|}) \gamma^N(r) |S^{N-1}(1)| dr|^2}{\int_0^\infty r^{N-1} \frac{F}{\gamma^{\otimes N}}(r \frac{x}{|x|}) \gamma^N(r) |S^{N-1}(1)| dr} \\ &\leq \int_{\mathbb{R}^N} \gamma^{\otimes N}(x) \int_{r=0}^{\infty} r^{N-1} \gamma^N(r) |S^{N-1}(1)| \frac{|\nabla [\frac{F}{\gamma^{\otimes N}}(r \frac{x}{|x|})]|^2}{\frac{F}{\gamma^{\otimes N}}(r \frac{x}{|x|})} dr dx. \end{aligned}$$

We simplify the last expression by introducing $\rho = F/\gamma^{\otimes N}$ and using the chain rule in Lemma 37 and the identity in Lemma 38. The end result is:

$$\begin{aligned} I(\check{F}|\gamma^{\otimes N}) &\leq \int_{\mathbb{R}^N} \gamma^{\otimes N}(x) \int_{r=0}^{\infty} r^{N-1} \frac{r^2}{|x|^2} \gamma^N(r) |S^{N-1}(1)| \\ &\quad \times \frac{|\nabla [\frac{F}{\gamma^{\otimes N}}](r \frac{x}{|x|})|^2 - ([\nabla \frac{F}{\gamma^{\otimes N}}](r \frac{x}{|x|}) \cdot \frac{x}{|x|})^2}{\frac{F}{\gamma^{\otimes N}}(r \frac{x}{|x|})} dr dx \\ &= \frac{1}{N-2} \int_{\mathbb{R}^N} |y|^2 \gamma^{\otimes N}(y) \frac{|\nabla \rho|^2(y) - (\nabla \rho(y) \cdot y)^2 / |y|^2}{\rho(y)} dy. \quad \square \end{aligned}$$

The proof of Proposition 36 will require the vanishing of the boundary terms arising from an integration by parts. The following lemma serves that purpose.

Lemma 39. *Let $\int_{\mathbb{R}} f(x) dx = 1$, $\int_{\mathbb{R}} x^2 f(x) dx = 1$ and $I(f|\gamma) < \infty$. Then $xf'(x) \in L^1(\mathbb{R})$ and $\lim_{x \rightarrow \pm\infty} xf(x) = 0$.*

Proof. We first recall that

$$\frac{[(f/\gamma)'(x)]^2}{(f/\gamma)(x)} \gamma(x) = \frac{(f'(x) + xf(x))^2}{f(x)}.$$

Thus $I(f|\gamma) = \int_{\mathbb{R}} \frac{(f'(x) + xf(x))^2}{f(x)} dx$. To show the integrability of $xf'(x)$ on \mathbb{R} , we write: for any $A > 0$,

$$\begin{aligned} \int_{\mathbb{R}} |xf'(x)| dx &\leq \int_{\mathbb{R}} |x[f'(x) + xf(x)]| dx + \int_{\mathbb{R}} x^2 f(x) dx \\ &= \int_{\mathbb{R}} |x| \sqrt{f(x)} \frac{|f'(x) + xf(x)|}{\sqrt{f(x)}} |dx + 1 \\ &\leq \sqrt{I(f|\gamma)} + 1. \end{aligned}$$

We now use integration by parts on $xf'(x)$ to show that $\lim_{A \rightarrow \infty} Af(A)$ exists, and deduce that $\lim_{A \rightarrow \infty} Af(A) = 0$. $\lim_{A \rightarrow \infty} A[f(A)] = \int_0^\infty xf'(x)dx + \int_0^\infty f(x)dx$. Thus $\lim_{A \rightarrow \infty} Af(A) = 0$ since $f(x)$ is integrable. The claim for $\lim_{A \rightarrow \infty} Af(-A) = 0$ is similarly obtained. \square

We are now ready to prove Proposition 36.

Proof of Proposition 36. Case 1: $N \geq 3$ We plug in $f^{\otimes N}$ for F in Proposition 35, then ρ can be expressed as

$$\rho(x) = \frac{f(x_1) \dots f(x_N)}{\gamma(x_1) \dots \gamma(x_N)} =: \rho_1^{\otimes N}.$$

Due to the product structure of ρ , it is convenient to write its gradient as

$$\nabla \rho(x_1, \dots, x_N) = \rho(x_1, \dots, x_N) \left(\frac{\rho'_1(x_1)}{\rho_1(x_1)}, \dots, \frac{\rho'_1(x_N)}{\rho_1(x_N)} \right).$$

We first consider the inequality (ii) in Proposition 35. In order to justify an integration by parts later on, we fix $A \gg 1$ and first take our integrals on the region $I_N := [-A, A]^N$. Afterwards, we can let $A \rightarrow \infty$. Expanding the sums $|x|^2 = x_1^2 + \dots + x_N^2$ and $|\nabla \rho|^2 = \rho(x)^2 \sum_{j=1}^N \left(\frac{\rho'_1(x_j)}{\rho_1(x_j)} \right)^2$, we have:

$$\begin{aligned} \int_{I_N} |x|^2 \gamma^{\otimes N}(x) \frac{|\nabla \rho(x)|^2}{\rho(x)} d^N x &= N \int_{-A}^A \gamma(x_1) x_1^2 \frac{\rho'_1(x_1)^2}{\rho_1(x_1)} dx_1 \left(\int_{-A}^A f(x_2) dx_2 \right)^{N-1} \\ &+ N(N-1) \int_{-A}^A x^2 f(x) dx \left(\int_{-A}^A f(x_2) dx_2 \right)^{N-2} \int_{-A}^A \gamma(x_1) \frac{\rho'_1(x_1)^2}{\rho_1(x_1)} dx_1 \end{aligned}$$

similarly, the second term in (ii) gives us the following.

$$\begin{aligned} - \int_{I_N} \gamma^{\otimes N}(x) \frac{(\nabla \rho(x) \cdot x)^2}{\rho(x)} dx &= -N \int_{-A}^A x_1^2 \gamma(x_1) \frac{\rho'_1(x_1)^2}{\rho_1(x_1)} dx_1 \left(\int_{-A}^A f(x_2) dx_2 \right)^{N-1} \\ &- N(N-1) \left(\int_{-A}^A \gamma(x_1) x_1 \rho'_1(x_1) \right)^2 \left(\int_{-A}^A f(x_2) dx_2 \right)^{N-2}. \end{aligned}$$

This last terms simplifies since an integration by parts, together with Lemma 39, shows that

$$- \int_{-A}^A \gamma(x_1) x_1 \rho'_1(x_1) dx_1 = A[f(A) + f(-A)] + \int_{-A}^A (\gamma(x_1) \rho_1(x_1) - x_1^2 \gamma(x_1) \rho_1(x_1)) dx_1.$$

We will set the above to be δ_A . Lemma 39 implies that $\lim_{A \rightarrow \infty} \delta_A = 1 - \int_{\mathbb{R}} x_1^2 f(x_1) dx_1 = 0$. Therefore,

$$\begin{aligned} I(\check{f}^N | \gamma^{\otimes N}) &\leq \lim_{A \rightarrow \infty} \frac{1}{N-2} \int_{I_N} \gamma^{\otimes N}(x) \frac{|x|^2 |\nabla \rho(x)|^2 - (\nabla \rho(x) \cdot x)^2}{\rho(x)} dx \\ &= \frac{N(N-1)}{(N-2)} \lim_{A \rightarrow \infty} \left[\int_{-A}^A x^2 f(x) dx \left(\int_{-A}^A f(x_2) dx_2 \right)^{N-2} \times \right. \\ &\quad \left. \int_{-A}^A \gamma(x_1) \frac{\rho'_1(x_1)^2}{\rho_1(x_1)} dx_1 - \delta_A^2 \left(\int_{-A}^A f(x_2) dx_2 \right)^{N-2} \right] \\ &= N \frac{N-1}{N-2} I(f | \gamma). \end{aligned}$$

Finally, from 1 in Proposition 35, the coefficient of $I(\hat{f}^N|\sigma^N)$ is simply $\frac{N-2}{N}$ times the coefficient of $I(\hat{f}^N|\gamma^{\otimes N})$ which is $N(1 - \frac{1}{N})$. This completes the proof for the case $N \geq 3$.

Case 2: $N = 2$

In this case, we cannot use Proposition 35. Instead, we treat with the rescaled state \hat{f}^2 directly. Note that we have the formula: $\hat{f}^2(\sqrt{2} \cos \theta, \sqrt{2} \sin \theta) = 2\pi \int_0^\infty r f(r \cos \theta) f(r \sin \theta) dr$, which, using $\rho(x) = f(x)/\gamma(x)$, can be rewritten as

$$\hat{f}^2(\sqrt{2} \cos \theta, \sqrt{2} \sin \theta) = \int_0^\infty r e^{-r^2/2} \rho(r \cos \theta) \rho(r \sin \theta) dr.$$

The relative Fisher information for \hat{f}^2 is given by $I(\hat{f}^2|\sigma^2) = \int_0^{2\pi} \frac{d\theta}{2\pi} \frac{(\frac{1}{\sqrt{2}} \frac{d}{d\theta} \hat{f}^2)^2}{\hat{f}^2}$. Since $r e^{-r^2/2}$ is a weight and the mapping $u \mapsto \frac{u'^2}{u}$ is convex, Jensen's inequality can be applied. This gives:

$$\begin{aligned} I(\hat{f}^2|\sigma^2) &= \frac{1}{4\pi} \int_0^{2\pi} d\theta \frac{(\frac{d}{d\theta} \hat{f}^2)^2}{\hat{f}^2} \leq \frac{1}{4\pi} \int_0^{2\pi} d\theta \int_{r=0}^\infty dr r e^{-r^2/2} \frac{(\frac{d}{d\theta} \rho(r \cos \theta) \rho(r \sin \theta))^2}{\rho(r \cos \theta) \rho(r \sin \theta)} \\ &= \frac{1}{4\pi} \int_0^{2\pi} d\theta \int_{r=0}^\infty dr r e^{-r^2/2} \left[2 \frac{\rho'(r \cos \theta)^2}{\rho(r \cos \theta)} \rho(r \sin \theta) (r \sin \theta)^2 \right. \\ &\quad \left. - 2 \rho'(r \cos \theta) \rho'(r \sin \theta) r \cos \theta r \sin \theta \right] \\ &= \int_{x \in \mathbb{R}} \gamma(x) \frac{\rho'(x)^2}{\rho(x)} dx \int_{y \in \mathbb{R}} \gamma(y) \rho(y) y^2 dy - \left(\int_{\mathbb{R}} \rho'(x) \gamma(x) x dx \right)^2 \\ &= I(f|\gamma), \end{aligned}$$

where we used the fact that $\int_{\mathbb{R}} f(y) y^2 dy = 1$, and that $\int_{\mathbb{R}} \rho'(x) \gamma(x) x dx = 0$ which follows from integration by parts. □

6 Conclusion and open questions

In this article, we considered the rescaled states \hat{f}^N , defined as the push-forward of $f^{\otimes N}$ by the mapping $x \mapsto \sqrt{N}x/|x| \in \mathcal{K}^N$, and established their chaoticity quantitatively, in the sense of Kac, using the Wasserstein metrics, in the sense of entropy, Fisher-information, and in the sense of L^1 for $f \in \text{ALip}(r) \subset L^1(\mathbb{R})$. The rescaled states provide sequences of measures on Kac's sphere which are chaotic to f when f is not necessarily in $L^1 \cap L^p$ for some $p > 1$.

Many interesting questions remain unanswered. A first set of questions concerns the set $\text{ALip}(r)$ and L^1 chaoticity. We recall from Lemma 20 that $\text{ALip}(r)$ is a subset of a Besov space. Can the L^1 chaoticity result in Theorem 10 be extended to all f in this Besov space? Also, can the space ALip be characterized further?

Another set of questions we left open concerns quantitative entropic chaoticity. The structure of the rescaled states allowed us to prove entropic chaoticity relatively easily. But in order to show quantitative entropic chaoticity, we used Theorem 28 (an improvement to [10, Theorem 4.17]), whose proof relies on a comparison with conditioned tensor products states $[f^{\otimes N}]_N$. This required f to have a finite relative Fisher information and that $f \in \mathcal{P}_{4+\epsilon}(\mathbb{R}) \cap L^p(\mathbb{R})$. It would be interesting to get rid of these restrictions. Also, avoiding the use of Theorem 28 would make our quantitative entropic chaos result independent of the conditioned tensor products. Moreover, since the rescaled tensor product \hat{f}^N requires less assumptions on f than $[f^{\otimes N}]_N$ does, one could adapt the proof of Theorem 28 in order to obtain quantitative entropic chaos rates valid for a broader class of sequences, by comparing them directly with \hat{f}^N .

We left open the problem of quantitative rates of Fisher-information chaos for the rescaled tensor product states.

Finally, we mention the following question. Do we have the limit:

$$\lim_{N \rightarrow \infty} H\left(f^{\otimes k} \mid \Pi_k \hat{f}^N\right) = 0$$

under some conditions? Stating the analogous question for $H(\Pi_k \hat{f}^N \mid f^{\otimes k})$ requires precaution. Even if $f \ll \gamma$, we can have $\Pi_k \hat{f}^N \not\ll f^{\otimes k}$. This occurs for example when $f([-a, a]) = 0$ for some $a > 0$, but f is not compactly supported.

7 Appendix

We now aim to prove Theorem 22. As in [3, Theorem 12], the proof uses the following result:

Lemma 40 (Legendre representation of the entropy). *Let E be a locally compact metric space, let $\mu, \nu \in \mathcal{P}(E)$. Then*

$$H(\mu \mid \nu) = \sup \left\{ \int \phi d\mu - \log \int e^\phi d\nu : \phi \in C_b(E) \text{ and Lipschitz} \right\}. \quad (7.1)$$

Moreover, one can restrict the supremum to functions satisfying $\int e^\phi d\nu = 1$.

Proof. Equation (7.1), without the Lipschitz condition, is part of the folklore; see for instance [12, Theorem B.2] for a proof in the case of E compact. From there, the restriction to Lipschitz functions is straightforward: since E is locally compact, one can restrict the supremum to continuous and compactly supported functions, which can further be approximated by Lipschitz functions in the supremum norm so that both integrals in (7.1) are close to the originals. We omit the details. \square

Proof of Theorem 22. We follow the proof of [3, Theorem 12]. Fix $\epsilon > 0$. From Lemma 40, there exists $\phi \in C_b(\mathbb{R})$ L -Lipschitz such that $\int_{\mathbb{R}} e^{\phi(x)} \gamma(x) dx = 1$ and

$$\int_{\mathbb{R}} \phi(x) f(dx) \geq H(f \mid \gamma) - \epsilon. \quad (7.2)$$

Define $\Phi \in C_b(\mathbb{R}^N)$ as $\Phi(x) = \phi(x_1) + \dots + \phi(x_N)$. Let $Z^N = (Z_1, \dots, Z_N) \sim \gamma^{\otimes N}$, thus $\hat{Z}^N \sim \sigma^N$. From (7.1) and symmetry, we obtain

$$\begin{aligned} \frac{1}{N} H(F^N \mid \sigma^N) &\geq \frac{1}{N} \int_{\mathcal{K}^N} \Phi(x) F^N(dx) - \frac{1}{N} \log \int_{\mathcal{K}^N} e^{\Phi(x)} \sigma^N(dx) \\ &= \int_{\mathcal{K}^N} \phi(x_1) F^N(dx) - \frac{1}{N} \log \mathbb{E}[e^{\Phi(\hat{Z}^N)}]. \end{aligned} \quad (7.3)$$

The idea is to replace $\Phi(\hat{Z}^N)$ by $\Phi(Z^N)$, so we now control their difference. Let $Q_N = \frac{1}{N} \sum_i Z_i^2$, thus $\hat{Z}_i^N = Q_N^{-1/2} Z_i$. Since ϕ is L -Lipschitz, the same argument that leads to (2.3) gives

$$|\Phi(\hat{Z}^N) - \Phi(Z^N)| \leq \sum_{i=1}^N |\phi(Q_N^{-1/2} Z_i) - \phi(Z_i)| \leq LN|Q_N - 1|.$$

Consider the event $A_N = \{|Q_N - 1| \leq N^{-1/2}\}$. Clearly,

$$\mathbb{E}[\mathbf{1}_{A_N} e^{\Phi(\hat{Z}^N)}] \leq \mathbb{E}[\mathbf{1}_{A_N} e^{\Phi(Z^N)} e^{LN|Q_N - 1|}] \leq e^{LN^{1/2}},$$

where we have used that $\mathbb{E}[e^{\Phi(Z^N)}] = 1$, because $\mathbb{E}[e^{\phi(Z_i)}] = 1$. Now, since \hat{Z}^N and Q_N are independent, we have $\mathbb{E}[\mathbf{1}_{A_N} e^{\Phi(\hat{Z}^N)}] = \mathbb{P}(A_N)\mathbb{E}[e^{\Phi(\hat{Z}^N)}]$, thus

$$\frac{1}{N} \log \mathbb{E}[e^{\Phi(\hat{Z}^N)}] = \frac{1}{N} \log \frac{\mathbb{E}[\mathbf{1}_{A_N} e^{\Phi(\hat{Z}^N)}]}{\mathbb{P}(A_N)} \leq \frac{L}{N^{1/2}} - \frac{1}{N} \log \mathbb{P}(A_N).$$

By the Central Limit Theorem applied to the sequence Z_1^2, Z_2^2, \dots , we know that $\mathbb{P}(A_N)$ converges to some strictly positive quantity. From (7.3), we thus have

$$\begin{aligned} \liminf_{N \rightarrow \infty} \frac{1}{N} H(F^N | \sigma^N) &\geq \liminf_{N \rightarrow \infty} \left\{ \int_{\mathcal{K}^N} \phi(x_1) F^N(dx) - \frac{L}{N^{1/2}} + \frac{1}{N} \log \mathbb{P}(A_N) \right\} \\ &= \int_{\mathbb{R}} \phi(x) f(dx), \end{aligned}$$

where we have used that $\Pi_1 F^N$ converges weakly to f . Using (7.2) and letting $\epsilon \rightarrow 0$, the conclusion follows. \square

Proof of Lemma 30. Here, we describe the proof of [10, Theorem 4.13] and state the adjustments needed to prove (4.7). We recall that for $f \in L^1(\mathbb{R})$, $[f^{\otimes N}]_N$ denotes the conditioned tensor product state defined in Def. 29. If $f \in L^p(\mathbb{R})$ for some $p > 1$ and $\int_{\mathbb{R}} |x|^k f(x) dx < \infty$ for some $k > 2$, then $\{[f^{\otimes N}(x)]_N\}_N$ is entropically chaotic to f . When $k \geq 4$, this was proven in [3] without any rates. When $k \geq 6$, entropic chaoticity with a rate of $N^{-1/2}$ was proven in [10, Theorem 4.13]. Finally, when $2 < k < 4$, entropic chaoticity of the family $\{[f^{\otimes N}(x)]_N\}$ was proven in [5] without rates of convergence. The quantitative result in [10, Theorem 4.13] can be extended to the case $k > 4$. We mention the technical changes necessary to relax the finite 6^{th} moment requirement.

- The Fourier based Lemma 4.8 in [10] can be modified to densities g having only M_{2+r} moments for some $r \in (0, 1]$. The new statement becomes:

Let $g \in \mathcal{P}_{2+r}(\mathbb{R}) \cap L^p(\mathbb{R})$ for some $p > 1$,

$\int_{\mathbb{R}} xg(x)dx = 0, \int_{\mathbb{R}} x \otimes xg(x) dx = \text{Id}, \int_{\mathbb{R}} |x|^{2+r}g(x) dx = M_{2+r}$. Then:

1. $\exists \delta > 0$ such that $|\xi| \leq \delta \rightarrow |\hat{g}(\xi)| \leq e^{-|\xi|^2/4}$.
2. $\forall \delta > 0, \exists k = k(M_{2+r}, r, p, \|g\|_p, \delta) \in (0, 1)$ such that

$$\sup_{|\xi| \geq \delta} |\hat{g}(\xi)| \leq k(\delta)$$

The first claim above follows from the idea that if g has a finite $2 + r$ moment, then

$$\hat{g}(\xi) = 1 - \frac{|\xi|^2}{8} + R(\xi), \quad |R(\xi)| \leq \frac{h_r}{(r+1)(r+2)} M_{2+r}(g) |\xi|^{2+r},$$

where $h_r = \sup_{\theta \neq 0} |\theta|^{-r} |e^{-i\theta} - 1|$.

The second claim follows from [3, Theorem 2.7(i)], and the observation that $\int g \log g < \infty$, whenever $g \in L^p$.

- Using the above result, the local central limit theorem in [10, Theorem 4.6] can be extended to $g \in \mathcal{P}_{2+r}(\mathbb{R}) \cap L^p(\mathbb{R})$, $r \in (0, 1], p \in (1, \infty]$. So that if $g_N(x)$ is the iterated and renormalized convolution

$$g_N(x) = \sqrt{N}g^{(*N)}(\sqrt{N}x),$$

then $\exists N = N(p)$ and $C_{BE} = C(p, r, M_{2+r}(g), \|g\|_p)$ such that

$$\forall N \geq N(p), \quad \|g_N - \gamma\|_{\infty} \leq \frac{C_{BE}}{N^{r/2}}.$$

This can be proved exactly as in [10, Theorem 4.6], with the additional observation that

$$\sup_{\xi \neq 0} \frac{|\hat{g}_N(\xi) - \hat{\gamma}_N(\xi)|}{|\xi|^{2+r}} < \infty.$$

- The above observations can pass on to [10, Theorem 4.9], with a remainder term $R_N(x)/N^{r/2}$, with $R_N \in L^\infty$.
- It remains to show that even when f has only $4 + s$ moments, the quantities $\theta_{N,1}$ defined by eq. (4.18) in [10] stay close to 1. Using $|e^{-x^2} - 1| \leq C_\alpha |x|^{2\alpha}$ for any $\alpha \in (0, 1]$, one can show that Eq. (4.20) in [10] can be replaced by

$$|\theta_{N,1}(v) - 1| \leq \frac{Cl^2}{N^{1/2}} + O(N^{-s/4}) + C_\alpha \frac{|v|^{4\alpha}}{N^\alpha} \mathbf{1}_{[N^u, \infty)}(|V|)$$

provided $\alpha(1 - 4u) = s/4$. Choosing $u = \frac{s}{s+2}$ and $\alpha = \frac{s(s+2)}{8}$ allows us to carry on the same decomposition of f used in the proof of Theorem 4.13 in [10] and to arrive at (4.7).

We emphasize that the condition $\int_{\mathbb{R}} vf(v) dv = 0$ mentioned is [10, Theorem 4.13] is not required. □

Sketch of the proof of Lemma 31. Starting with the formula

$$I([f^{\otimes N}]_N | \sigma^N) = \int_{\mathcal{K}^N} |\nabla_S \log h_N(w)| [f^{\otimes N}]_N(dw)$$

where h_N is the density of $[f^{\otimes N}]_N$ with respect to σ^N , given by

$$\frac{d[f^{\otimes N}]_N}{d\sigma^N}(w) = \frac{f^{\otimes N}(w)}{\int_{\mathcal{K}^N} f^{\otimes N}(y) \sigma^N(dy)}.$$

We note that we can replace $h_N(w)$ by $\frac{f^{\otimes N}(w)}{\gamma^{\otimes N}(w)}$ since $\gamma^{\otimes N}$ is constant on \mathcal{K}^N .

Using $|\nabla_S G(v)|^2 \leq |\nabla G(v)|^2$ when G is defined on \mathbb{R}^N , we obtain (see Eq. (4.24) in [10]):

$$\frac{1}{N} I([f^{\otimes N}]_N | \sigma^N) \leq I(f|\gamma) + \int_{\mathbb{R}} \left| \frac{\nabla f(v)}{f(v)} + v \right|^2 (\theta_{N,1}(v) - 1) f(v) dv,$$

where the product $\theta_{N,1}(v)f(v)$ equals the first marginal of $[f^{\otimes N}]_N$ and, $\theta_{N,1}(v)$ satisfies $|\theta_{N,1}(v)| \leq C$ uniformly in N (see [10, Equation (4.20)]). Thus, $\frac{1}{N} I([f^{\otimes N}]_N | \sigma^N) \leq (1 + C)I(f|\gamma)$. □

References

- [1] F. Barthe, D. Cordero-Erausquin, and B. Maurey, *Entropy of spherical marginals and related inequalities*, J. Math. Pures Appl. (9) **86** (2006), no. 2, 89–99. MR2247452
- [2] E. A. Carlen, E. H. Lieb, and M. Loss, *A sharp analog of Young's inequality on S^N and related entropy inequalities*, J. Geom. Anal. **14** (2004), no. 3, 487–520. MR2077162
- [3] Eric A. Carlen, Maria C. Carvalho, Jonathan Le Roux, Michael Loss, and Cédric Villani, *Entropy and chaos in the Kac model*, Kinet. Relat. Models **3** (2010), no. 1, 85–122. MR2580955
- [4] Kleber Carrapatoso, *Quantitative and qualitative Kac's chaos on the Boltzmann's sphere*, Ann. Inst. Henri Poincaré Probab. Stat. **51** (2015), no. 3, 993–1039. MR3365971
- [5] Kleber Carrapatoso and Amit Einav, *Chaos and entropic chaos in Kac's model without high moments*, Electron. J. Probab. **18** (2013), no. 78, 38. MR3101644
- [6] Roberto Cortez, *Uniform propagation of chaos for Kac's 1D particle system*, J. Stat. Phys. **165** (2016), no. 6, 1102–1113. MR3575639

- [7] Roberto Cortez and Joaquin Fontbona, *Quantitative uniform propagation of chaos for Maxwell molecules*, Comm. Math. Phys. **357** (2018), no. 3, 913–941. MR3769742
- [8] Roberto Cortez and Hagop Tossounian, *On a thermostated Kac model with rescaling*, Ann. Henri Poincaré **22** (2021), no. 5, 1629–1668. MR4250744
- [9] Nicolas Fournier and Arnaud Guillin, *From a Kac-like particle system to the Landau equation for hard potentials and Maxwell molecules*, Ann. Sci. Éc. Norm. Supér. (4) **50** (2017), no. 1, 157–199. MR3621429
- [10] Maxime Hauray and Stéphane Mischler, *On Kac's chaos and related problems*, J. Funct. Anal. **266** (2014), no. 10, 6055–6157. MR3188710
- [11] M. Kac, *Foundations of kinetic theory*, Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, vol. III (Berkeley and Los Angeles), University of California Press, 1956, pp. 171–197. MR0084985
- [12] John Lott and Cédric Villani, *Ricci curvature for metric-measure spaces via optimal transport*, Ann. of Math. (2) **169** (2009), no. 3, 903–991. MR2480619
- [13] Stéphane Mischler and Clément Mouhot, *Kac's program in kinetic theory*, Invent. Math. **193** (2013), no. 1, 1–147. MR3069113
- [14] F. Otto and C. Villani, *Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality*, J. Funct. Anal. **173** (2000), no. 2, 361–400. MR1760620
- [15] Alain-Sol Sznitman, *Topics in propagation of chaos*, École d'Été de Probabilités de Saint-Flour XIX—1989, Lecture Notes in Math., vol. 1464, Springer, Berlin, 1991, pp. 165–251. MR1108185
- [16] Cédric Villani, *Optimal transport, old and new*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 338, Springer-Verlag, Berlin, 2009. MR2459454
- [17] Bengt von Bahr and Carl-Gustav Esseen, *Inequalities for the r th absolute moment of a sum of random variables, $1 \leq r \leq 2$* , Ann. Math. Statist. **36** (1965), 299–303. MR170407

Acknowledgments. We thank the two anonymous referees who provided insightful comments and suggestions that allowed us to improve the presentation of this article.

Electronic Journal of Probability

Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards
- Free for authors, free for readers
- Quick publication (no backlog)
- Secure publication (LOCKSS¹)
- Easy interface (EJMS²)

Economical model of EJP-ECP

- Non profit, sponsored by IMS³, BS⁴, ProjectEuclid⁵
- Purely electronic

Help keep the journal free and vigorous

- Donate to the IMS open access fund⁶ (click here to donate!)
- Submit your best articles to EJP-ECP
- Choose EJP-ECP over for-profit journals

¹LOCKSS: Lots of Copies Keep Stuff Safe <http://www.lockss.org/>

²EJMS: Electronic Journal Management System: <https://vtex.lt/services/ejms-peer-review/>

³IMS: Institute of Mathematical Statistics <http://www.imstat.org/>

⁴BS: Bernoulli Society <http://www.bernoulli-society.org/>

⁵Project Euclid: <https://projecteuclid.org/>

⁶IMS Open Access Fund: <https://imstat.org/shop/donation/>