

ENTROPY AND CHAOS IN THE KAC MODEL

Dedicated to Professor Giuseppe Toscani on the occasion of his sixtieth birthday

ERIC A. CARLEN

Department of Mathematics, Hill Center
Rutgers University
Piscataway, NJ 08854, U.S.A.

MARIA C. CARVALHO

Department of Mathematics and CMAF
University of Lisbon
1649-003 Lisbon, PORTUGAL

JONATHAN LE ROUX

Department of Information Physics and Computing
The University of Tokyo
7-3-1, Hongo, Bunkyo-ku, Tokyo 113-8656, JAPAN

MICHAEL LOSS

School of Mathematics
Georgia Institute of Technology
Atlanta GA, 30332, U.S.A.

CÉDRIC VILLANI

UMPA, ENS Lyon
University of Lyon
46 allée d'Italie, 69364 Lyon Cedex 07, FRANCE

(Communicated by the associate editor name)

ABSTRACT. We investigate the behavior in N of the N -particle entropy functional for Kac's stochastic model of Boltzmann dynamics, and its relation to the entropy function for solutions of Kac's one dimensional nonlinear model Boltzmann equation. We prove results that bring together the notion of propagation of chaos, which Kac introduced in the context of this model, with the problem of estimating the rate of equilibration in the model in entropic terms, showing that the entropic rate of convergence can be arbitrarily slow. Results proved here show that one can in fact use entropy production bounds in Kac's stochastic model to obtain entropic convergence bounds for his non linear model Boltzmann equation, though the problem of obtaining optimal lower bounds of this sort for the original Kac model remains open and the upper bounds obtained here show that this problem is somewhat subtle.

2000 *Mathematics Subject Classification.* Primary: 76P05, 60G50; Secondary: 54C70.

Key words and phrases. Entropy, propagation of chaos.

The work of E.C. and M.L. was partially supported U.S. National Science Foundation grant DMS 06-00037. The work of M.C.C. was partially supported by POCI/MAT/61931/2004.

1. Introduction.

1.1. The origins of the problem to be considered. In a remarkable paper [18] of 1956, Mark Kac investigated the probabilistic foundations of kinetic theory, and defined the notion of **propagation of chaos**, which has since then developed into an active field of probability.

Kac introduced the concept of propagation of chaos in connection with a specific stochastic process modeling binary collisions in a gas made of a large number N of identical molecules, and he was particularly concerned with its rate of equilibration, i.e., of its approach to stationarity.

While his ideas concerning propagation of chaos had an immediate resonance and impact, this was not the case with the issues he raised concerning rates of equilibration. These had to wait much longer for progress and development, as we shall relate below. In this paper, we bring these two lines of investigation back together, proving several theorems relating chaos and equilibration for the Kac walk.

1.2. The Kac walk. We begin with a precise description of the **Kac walk** as a model for the evolution of the distribution of velocities in a gas of like molecules undergoing binary collisions. For simplicity, Kac assumed the gas to be spatially homogeneous, and the velocities v_j ($1 \leq j \leq N$) to be one-dimensional. The latter assumption is incompatible with the conservation of both momentum and kinetic energy, so Kac only assumed conservation of the kinetic energy E , where

$$E = \frac{m}{2} \sum_{j=1}^N v_j^2,$$

with m denoting the mass of the particle species, and v_j denoting the velocity of the j th particle.

The natural state space for this system (i.e., state space for the walk) is the sphere $S^{N-1}(\sqrt{(2/m)E}) \subset \mathbb{R}^N$, the $(N-1)$ -dimensional sphere with radius $\sqrt{(2/m)E}$. For the sequel of the discussion, let us choose units in which the mass of each particle is 2. We also choose units in which the total value of the kinetic energy is N , so that the state space is $S^{N-1}(\sqrt{N})$, and each particle has unit mean kinetic energy. Let $V = (v_1, \dots, v_N)$ denote a generic point in $S^{N-1}(\sqrt{N})$.

Here is how to take a step of the Kac walk: First, randomly pick a pair (i, j) of distinct indices in $\{1, \dots, n\}$ uniformly from among all such pairs. The molecules i and j are the molecules that will “collide”. Second, pick a random angle θ uniformly from $[0, 2\pi)$. Then update $V = (v_1, \dots, v_N)$ by leaving v_k unchanged for $k \neq i, j$, and updating velocities v_i and v_j by rotating in the v_i, v_j plane as follows:

$$(v_i, v_j) \rightarrow \left((\cos \theta)v_i - (\sin \theta)v_j, (\sin \theta)v_i + (\cos \theta)v_j \right).$$

Let $R_{i,j,\theta}V$ denote the new point in $S^{N-1}(\sqrt{N})$ obtained in this way. This process, repeated again and again, is the **Kac walk** on $S^{N-1}(\sqrt{N})$.

Associated to the steps of this walk is the Markov transition operator Q_N on $L^2(S^{N-1}(\sqrt{N}), d\sigma^N)$ where σ^N is the uniform probability measure on $S^{N-1}(\sqrt{N})$. (This notation shall be used throughout the paper.)

If V_j denotes the position after the j th step of the walk, and φ is any continuous function on $S^{N-1}(\sqrt{N})$, the transition operator Q_N is defined by

$$Q_N \varphi(V) = \mathbb{E} \{ \varphi(V_{j+1}) \mid V_j = V \} .$$

From the description provided above, one finds that

$$Q_N \varphi(V) = \binom{N}{2}^{-1} \sum_{i < j} \frac{1}{2\pi} \int_{[0, 2\pi)} \varphi(R_{i,j,\theta} V) d\theta .$$

It is easily seen that σ^N is the unique invariant measure.

A closer match with the physics being modeled is attained if the steps of the walk arrive not in a metronome beat, but in a Poisson stream with the mean wait between steps being $1/N$. This ‘‘Poissonification’’ of the Kac walk yields a continuous time process on $S^{N-1}(\sqrt{N})$. Since Q_N is self adjoint on $L^2(S^{N-1}(\sqrt{N}), d\sigma^N)$, this process is reversible, and so if the law μ_0 of the initial state V_0 has a density F_0^N with respect to σ^N , then for all $t > 0$, the law μ_t of V_t has a density F_t^N with respect to σ^N , and F_t^N is the solution to the Cauchy problem

$$\frac{\partial}{\partial t} F^N = L_N F^N \quad \text{with} \quad \lim_{t \rightarrow 0} F_t^N = F_0^N \quad (1)$$

where $L_N = N(Q_N - I)$, and I is the identity operator. This equation is known as the **Kac master equation**, which is nothing other than the Kolmogorov forward equation for the continuous time Kac walk. The solution is of course given by

$$F_t = e^{tL_N} F_0 . \quad (2)$$

Since $V \rightarrow R_{i,j,\theta}$ is a rotation, it follows that for each positive integer k , Q_N preserves the subspace of $L^2(S^{N-1}(\sqrt{N}), d\sigma^N)$ consisting of spherical harmonics of degree no greater than k . Hence, all of the eigenfunctions of Q_N are spherical harmonics. Since the constant is the only spherical harmonic that is invariant under rotations, 1 is an eigenvalue of Q of multiplicity one.

Therefore, for any initial data F_0 in $L^2(S^{N-1}(\sqrt{N}), d\sigma^N)$, the solution $F_t^N = e^{L_N t} F_0^N$ of the Kac master equation satisfies

$$\lim_{t \rightarrow \infty} F_t^N = 1 . \quad (3)$$

We refer to the invariant density 1 as the *equilibrium*, and the process of approaching this limit as *equilibration*.

The rate at which this limit is achieved is physically interesting for reasons that will be explained shortly. But apart from its physical motivation, the problem is quite interesting on purely probabilistic grounds: While the subject of quantifying the rate of equilibration for random walks on large discrete sets has been vigorously developed in recent years, much less has been done in the case of continuous state spaces of high dimension, and the Kac walk is a very natural example.

Kac proposed to investigate the rate of equilibration for his walk in L^2 terms through the *spectral gap* of L_N : Define

$$\Delta_N = \sup \{ - \langle \varphi, L_N \varphi \rangle : \langle \varphi, 1 \rangle = 0 \quad \text{and} \quad \langle \varphi, \varphi \rangle = 1 \}$$

where the inner products are taken in $L^2(S^{N-1}(\sqrt{N}), d\sigma^N)$. In his paper [18], Kac conjectured that $\liminf_{N \rightarrow \infty} \Delta_N > 0$.

Since one already knows that the eigenfunctions of L_N are spherical harmonics, this may seem like a trivial problem. In fact, it is very easy to guess the exact value for Δ_N and the corresponding eigenfunction. Indeed, it is natural to suppose that the eigenfunction must be a simple symmetric, even polynomial in the velocities v_j . The simplest such thing, $\sum_{j=1}^N v_j^2$, is simply a constant on $S^{N-1}(\sqrt{N})$, so one

might try

$$\varphi_{\text{gap}} = \sum_{j=1}^N (v_j^4 - \langle v_j^4, 1 \rangle).$$

(The constant being subtracted to ensure orthogonality to 1 can be easily computed; see [8], and this is indeed a spherical harmonic.) Physical reasoning, based on linearizing the Boltzmann-Kac equation to be discussed shortly, gives further evidence that φ_{gap} should in fact be the gap eigenfunction. Using this as a trial function, one readily computes what should be — and does turn out to be — the value of Δ_N :

$$\Delta_N = \frac{1}{2} \frac{N+2}{N-1}. \quad (4)$$

However, while one can explicitly compute as many eigenvalues as one wants to, there is no monotonicity argument to rule out the proposition that the gap eigenvalue might come from a spherical harmonic of large degree.

Kac's conjecture that $\liminf_{N \rightarrow \infty} \Delta_N > 0$ was first proved by Janvresse [17]. Her method did not yield the exact value for Δ_N . The first proof that (4) is actually correct was given in [8]; see also [22] for a different approach. For a treatment of related problems, including physical three-dimensional momentum preserving collisions, see [9] and [11]. Another interesting approach to these problems has been developed recently by Caputo [5]. His method give the exact spectral gap in one dimension very simply, and though it does not give the exact gap in three dimensions, it does at least provide a uniform bound away from zero.

These results enable us to quantify (3) as follows:

$$\|F_t^N - 1\|_{L^2(S^{N-1}(\sqrt{N}), d\sigma^N)} \leq e^{-t/2} \|F_0^N - 1\|_{L^2(S^{N-1}(\sqrt{N}), d\sigma^N)}$$

for all N and t . While the exponent is uniform in N , the shortcoming of this result will be familiar to many probabilists who have worked on rates of equilibration: *For natural sequences of initial data $\{F_0^N\}_{N \in \mathbb{N}}$, it will be the case that*

$$\|F_0^N\|_{L^2(S^{N-1}(\sqrt{N}), d\sigma^N)} \geq C^N$$

for some $C > 1$. Therefore, one still has to wait a *time proportional to N* before the bound starts providing evidence of equilibration.

Even worse, the badly behaved sequences of initial data mentioned above are exactly the ones of primary physical interest — the *chaotic* sequences, in which for large N the coordinate functions v_j are “nearly independent and identically distributed” under the law $\mu^{(N)} = F_0^N \sigma^N$.

1.3. Kac's notion of chaos. To state the precise definition, we first introduce some notation that will be used throughout the paper: Given any probability measure $\mu^{(N)}$ on $S^{N-1}(\sqrt{N})$, and any positive integer $k < N$, let $P_k(\mu^{(N)})$ denote the marginal measure of $\mu^{(N)}$ for the first k velocities. In formulas: whenever A is a Borel subset of \mathbb{R}^k ,

$$P_k(\mu^{(N)})[A] = \mu^{(N)}[\{(v_1, \dots, v_k) \in A\}].$$

In the sequel, we only consider symmetric measures, so there is nothing particular in considering the *first* k velocities. Chaos means that $P_k \mu^{(N)}$ is well approximated by $\mu^{\otimes k}$, a distribution of k *independent* particles when N is large. Here is a more precise definition:

Definition 1 (chaos). Let μ be a given Borel probability measure on \mathbb{R} . For each positive integer N , let $\mu^{(N)}$ be a probability measure on $S^{N-1}(\sqrt{N})$. Then the sequence of probability measures $\{\mu^{(N)}\}_{N \in \mathbb{N}}$ is said to be μ -**chaotic** in case:

- (i) Each $\mu^{(N)}$ is symmetric under interchange of the variables v_1, \dots, v_N ;
- (ii) For each fixed positive integer k , the marginal $P_k \mu^{(N)}$ of $\mu^{(N)}$ (marginal on the first k velocities) converges to the k -fold tensor product $\mu^{\otimes k}$, as $N \rightarrow \infty$, in the sense of weak convergence against bounded continuous test functions. That is, whenever $\chi(v_1, \dots, v_k)$ is a bounded continuous test function of k variables, then

$$\int \chi(v_1, \dots, v_k) d\mu^{(N)}(v_1, \dots, v_N) \xrightarrow{N \rightarrow \infty} \int \chi(v_1, \dots, v_k) d\mu(v_1) \cdots d\mu(v_k). \quad (5)$$

Property (ii) says that $\mu^{(N)}$ is well approximated by $(P_1 \mu^{(N)})^{\otimes N}$ as $N \rightarrow \infty$, in the weak sense of convergence against test functions depending on a finite number of variables.

Besides being archetypal, the following well-known example will play an important role in this paper. It has quite an ancient history, going back —at least— to Mehler [23] in 1866. For a more recent reference, see [24].

Example 2. Let

$$\gamma(v) = \frac{e^{-v^2/2}}{\sqrt{2\pi}}. \quad (6)$$

Then, $\{\sigma^N\}_{N \in \mathbb{N}}$ is $\gamma(v)dv$ chaotic. Indeed, this follows easily from the explicit computation

$$P_k \sigma^N = \left(1 - \frac{s^2}{N}\right)^{\frac{N-k-2}{2}} \frac{|S^{N-k-1}|}{N^{k/2} |S^{N-1}|} \mathcal{L}_k, \quad \text{where } |S^{N-1}| = \frac{2\pi^{N/2}}{\Gamma(\frac{N}{2})}, \quad (7)$$

and where \mathcal{L}_k is the k -dimensional Lebesgue measure.

Now, let f be some probability density on \mathbb{R} , and (with the same notation as in the above example) suppose that $\{F^N \sigma^N\}_{N \in \mathbb{N}}$ is an $f(v)dv$ chaotic family. For each N , let $F^N(t, \cdot)$ denote the solution of (1) at time t , starting from the initial data F^N . The main result that Kac did prove in [18] is that for each $t > 0$, $\{F^N(t, \cdot) \sigma^N\}_{N \in \mathbb{N}}$ is still a chaotic family; this property is referred to as **propagation of chaos**. Indeed, $\{F^N(t, \cdot) \sigma^N\}_{N \in \mathbb{N}}$ is $f(t, v)dv$ chaotic, where $f(t, v)$ is the solution of the following Cauchy problem:

$$\begin{cases} f(0, \cdot) = f; \\ \frac{\partial f}{\partial t}(t, v) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \int_{\mathbb{R}} [f(v', t) f(v'_*, t) - f(v, t) f(v_*, t)] dv_* d\theta, \end{cases} \quad (8)$$

and

$$v' = (\cos \theta) v - (\sin \theta) v_*, \quad v'_* = (\sin \theta) v + (\cos \theta) v_*.$$

This nonlinear equation is a model Boltzmann equation, which we shall call the **Boltzmann-Kac equation** (as opposed to the Kac master equation). The quadratic nonlinearity on the right is a reflection of the fact that Q_N models a binary collision process, and of Kac's notion of chaos: Indeed, the time derivative of $P_1(e^{tL_N} F^N)$ may be expressed in terms of a linear operation on $P_2 F^N$, and then this is well approximated by the tensor product $f \otimes f$ in the limit $N \rightarrow \infty$.

The program Kac set forth in [18] was to investigate the behavior of solutions of (8) in terms of the behavior of solutions of the the Kac master equation (1). In particular, concerning equilibration,

$$\lim_{t \rightarrow \infty} F_t^N = 1 \Rightarrow \lim_{t \rightarrow \infty} P_1(F_t^N \sigma^N) = \frac{|S^N - 2|}{N^{1/2}|S^{N-1}|} \left(1 - \frac{v_1^2}{N}\right)^{(N-3)/2} \approx \gamma(v_1)$$

for large N , and thus Kac's theorem can be used to relate the rate of equilibration in the Kac master equation to the rate of convergence in

$$\lim_{t \rightarrow \infty} f(v, t) = \gamma(v)$$

for solutions $f(v, t)$ of (8). Once this would be carried out, one would then like to do the same for the actual Boltzmann equation for three dimensional velocities with conservation of both energy and momentum.

As we have indicated, an L^2 analysis of the rate of equilibration for solutions of the Kac master equation does not shed much light on the large time behavior of solutions of (8). What *would* do this is very natural in the context of the Boltzmann equation: *an entropy production estimate*.

1.4. Convergence to equilibrium and entropy inequalities. If μ and ν are two probability measures on a measurable space \mathcal{X} , their relative entropy is defined by the formula

$$H(\mu|\nu) = \int h \log h \, d\nu \quad h = \frac{d\mu}{d\nu},$$

with the understanding that $H(\mu|\nu) = +\infty$ if μ is not absolutely continuous with respect to ν . In particular,

- if f is a probability density on \mathbb{R} , then its relative entropy with respect to γ (identified with a probability measure) is

$$H(f|\gamma) = \int_{\mathbb{R}} f(v) \log \frac{f(v)}{\gamma(v)} \, dv;$$

- if F^N is a probability density on $S^{N-1}(\sqrt{N})$, then its relative entropy with respect to the uniform probability measure σ^N is

$$H_N(F^N) := H(F^N \sigma^N | \sigma^N) = \int_{S^{N-1}(\sqrt{N})} F^N(v) \log F^N(v) \, d\sigma^N(v).$$

The well-known Csiszar-Kullback-Leibler-Pinsker inequality states that

$$H(\mu|\nu) \geq \|\mu - \nu\|_{TV}^2/2, \tag{9}$$

where the subscript “TV” stands for the total variation norm.

So the relative entropy measures a deviation from equilibrium, just like the L^2 norm, and it is natural to try to quantify the rate of equilibration for the Kac master equation by studying $H_N(F_t^N)$ for solutions: If F_t^N is a solution,

$$\frac{d}{dt} H_N(F_t^N) = \int_{S^{N-1}(\sqrt{N})} \log(F_t^N) L_N F_t^N \, d\sigma^N = \langle \log(F_t^N), L_N F_t^N \rangle.$$

In analogy with the definition of the spectral gap Δ_N , define the *entropy production constant* Γ_N by

$$\Gamma_N = \inf \frac{-\langle \log(F^N), L_N F^N \rangle}{H_N(F^N)}$$

where the infimum is taken over all probability densities F^N on $S^{N-1}(\sqrt{N})$ with $H_N(F^N) < \infty$.

The entropic analog of the Kac conjecture would be that there exists a $c > 0$ with $\Gamma_N \geq c$ for all N . This would imply that

$$H_N(F_t^N) \leq e^{-ct} H_N(F_0^N), \quad (10)$$

and hence that

$$\|F_t^N \sigma^N - \sigma^N\|_{TV}^2 \leq 2e^{-ct} H_N(F_0^N).$$

There is an absolutely crucial difference between this and the L^2 estimate that we obtained earlier, and it lies in the *extensivity of the entropy*. Suppose that $\{F_0^N\}_{N \in \mathbb{N}}$ is an $f_0(v)dv$ -chaotic family of densities on $S^{N-1}(\sqrt{N})$. Then according to Kac's theorem, $\{F_t^N\}_{N \in \mathbb{N}}$ is an $f(v, t)dv$ -chaotic family of densities on $S^{N-1}(\sqrt{N})$, where $f(v, t)$ is the solution of the Boltzmann–Kac equation with initial data $f_0(v)$. Because of the near product structure of F_N^t , one might expect that for each t , and large N ,

$$H_N(F_t^N) \approx NH(f(t, \cdot)dv|\gamma dv). \quad (11)$$

It is the proportionality to N that we refer to as extensivity. Since this factor of N would appear on both sides of (10) if we substituted (11) in on both sides, we can cancel off the N , and obtain, in the large N limit

$$H(f(t, \cdot)dv|\gamma dv) \leq e^{-ct} H(f_0 dv|\gamma dv).$$

We could now apply (9) to this and conclude that for solutions $f(v, t)$ of the Boltzmann–Kac equation,

$$\|f(\cdot, t)dv - \gamma dv\|_{TV}^2 \leq 2e^{-ct} H(f_0 dv|\gamma dv).$$

Such a bound would be very desirable to have for the Boltzmann–Kac equation, and this motivates the enquiry into the exact behavior of the entropy production constant Γ_N .

It turns out that estimating the entropy production constant Γ_N is a much more subtle problem than that of estimating the spectral gap Δ_N . Unfortunately, the best information that is known at present is

$$\Gamma_N \geq \frac{2}{N-1}.$$

There are two different proofs of this result. The first, due to Villani, can be found under Theorem 6.1 in [26]. (The bound $2/(N-1)$ is what one gets from the argument in [26] making some simplifications that are admissible in the special case of the original Kac model considered here.) The second, due to Carlen and Loss, is an entropic adaptation of the argument used in [8] to determine the spectral gap. It can be found under Lemma 2.4 in [7] using Theorem 2.5 there. It was conjectured in [26] that these bounds are essentially sharp; i.e., that

$$\Gamma_N = \mathcal{O}\left(\frac{1}{N}\right).$$

However, this is not so clear at present. In fact, it had remained an open problem whether there was even a sequence $\{F^N\}$ of densities for which

$$\lim_{N \rightarrow \infty} \frac{-\langle \log(F^N), L_N F^N \rangle}{H_N(F^N)} = 0 \quad (12)$$

with convergence at any rate at all. The following theorem settles this issue:

Theorem 3. For each $c > 0$, there is a probability density f on \mathbb{R} with $\int_{\mathbb{R}} vf(v)dv = 0$ and $\int_{\mathbb{R}} v^2 f(v)dv = 1$, and an f -chaotic family $\{F^N \sigma^N\}_{N \in \mathbb{N}}$ such that

$$\limsup_{N \rightarrow \infty} \frac{-\langle \log(F^N), L_N F^N \rangle}{H_N(F^N)} \leq c .$$

For each c , the density f is smooth and bounded, and has moments of all orders.

Once one has this, an easy diagonal argument produces a sequence $\{F^N\}_{N \in \mathbb{N}}$ satisfying (12).

While this would seem to be bad news for Kac's program, it only shows that one cannot have a *universal* bound on the ratio defining Γ_N , valid for *all* probability densities F_N on $S^{N-1}(\sqrt{N})$. Theorem 3 does not rule out the possibility that there is a *conditional* bound on this ratio, holding for all F^N in an f -chaotic family with some condition on f .

Indeed, we shall see that the densities f used to prove Theorem 3 have a fourth moment that diverges as c tends to zero. As far as we now know, Theorem 3 might become false under the additional assumption of a fixed bound on the fourth moment of f . This would be very interesting since bounds on the fourth moments are well known to be preserved by solutions of the Boltzmann–Kac equation, so such a condition on the initial data would propagate.

Moreover, it is known [12] that even for smooth initial data f with $\int_{\mathbb{R}} vf(v)dv = 0$ and $\int_{\mathbb{R}} v^2 f(v)dv = 1$, solutions $f(v, t)$ of the Boltzmann–Kac equation can have $\|f(\cdot, t) - \gamma\|_{L^1(\mathbb{R})}$ approach zero arbitrarily slowly – for example, like

$$\frac{1}{1 + \log(1 + \log(1 + \log(1 + t)))} ,$$

or the same thing with as many logarithms as one might wish. This however can happen only when the density f has very long tails so that $\int_{\mathbb{R}} v^2 f(v)dv$ just barely converges. A bound on the fourth moment, which would ensure good behavior of the tails is therefore a plausible condition to impose if one seeks a lower bound on the rate of convergence.

Finally, if one modifies the Kac walk so that pairs of molecules i, j with high values of $v_i^2 + v_j^2$ run *much* faster, then one can prove a uniform positive lower bound on Γ_N ; see [26, Section 6]. Thus, Theorem 3 displays the subtleties that beset Kac's program, but it does not by any means terminate it. In fact it raises a very interesting question: What sort of conditional bound on Γ_N might hold for the Kac model? But we shall not come to that in this paper; there are more basic issues to be settled first.

1.5. Conditioned tensor products. The proof of Theorem 3 naturally requires the *construction of chaotic data*, and this raises the following question:

Question 1. Let f be a probability density on \mathbb{R} with

$$\int_{\mathbb{R}} vf(v)dv = 0 \quad , \quad \int_{\mathbb{R}} v^2 f(v)dv = 1 \quad , \quad (13)$$

and finite entropy. Is it true that there is an $f(v)dv$ -chaotic family of densities $\{F^N\}_{N \in \mathbb{N}}$ on $S^{N-1}(\sqrt{N})$?

Question 1 may seem trivial at first sight, and actually was treated by Kac in a rather cavalier fashion. Indeed, there is an obvious procedure for generating chaotic initial data, which may be described as follows.

Suppose that $\mu(dv)$ is a probability measure on \mathbb{R} satisfying (13). Consider the tensor product measure $\mu^{\otimes N}$ and *condition* (restrict) it to the sphere $S^{N-1}(\sqrt{N})$. By the law of large numbers, $\sum_{j=1}^N v_j^2 \approx N$ for large N , almost surely with respect to $\mu^{\otimes N}$, so this measure is roughly concentrated on $S^{N-1}(\sqrt{N})$, and the conditioning should not modify it too much.

An important instance where this is obviously true is the particular case when $\mu = \gamma$: Then F^N is just the uniform measure σ^N , and the explicit formula (7) certainly guarantees that F^N is γ -chaotic in a very strong sense.

But for more general data, the extent to which $\mu^{\otimes N}$ is actually concentrated on $S^{N-1}(\sqrt{N})$ is not so obvious. Assume that μ has a density f , so $f^{\otimes N}$ is the density of $\mu^{\otimes N}$; then the restriction of $f^{\otimes N}$ to $S^{N-1}(\sqrt{N})$ (which is a set of zero measure) might just *not* be well-defined under the conditions (13) alone. Whether or not this is the case depends on the fluctuations of $\sum_{j=1}^N v_j^2$ about N ; i.e., on how well $\mu^{\otimes N}$ is concentrated on $S^{N-1}(\sqrt{N})$, as measured by the variance of v^2 with respect to $f(v)dv$. Again, *this will be governed by a fourth moment condition*.

In what follows we shall use the following notation: For a probability density $f(v)$ on \mathbb{R} , satisfying $\int f(v)v^2 dv = 1$, Let Σ^2 denote the variance of v^2 under $f(v)dv$:

$$\Sigma := \sqrt{\int_{\mathbb{R}} (v^2 - 1)^2 f(v) dv}.$$

Also, define

$$Z_N(f, r) := \int_{S^{N-1}(r)} f^{\otimes N} d\sigma_r^N,$$

where $S^{N-1}(r)$ is the sphere of radius r in \mathbb{R}^N , and σ_r^N is the uniform probability measure on that sphere.

The technical core of our results lies in the following estimates, that can be seen as a version of the **local central limit theorem**.

Theorem 4 (Estimates on a conditioned tensor product). With the above notation and under assumptions (13) and

$$\int_{\mathbb{R}} v^4 f(v) dv < +\infty \quad \int_{\mathbb{R}} f^p < +\infty \quad (14)$$

for some $p > 1$,

$$Z_N(f, r) = \gamma^{(N)}(r) \frac{\sqrt{2}}{\Sigma} \frac{\alpha_N(N)}{\alpha_N(r^2)} \left(e^{-\frac{(r^2-N)^2}{2N\Sigma^2}} + \varepsilon(N, f, r) \right), \quad (15)$$

where $\gamma^{(N)}(r)$ is the restriction of $\gamma^{\otimes N}$ to $S^{N-1}(r)$,

$$\alpha_N(u) = u^{\frac{N}{2}-1} e^{-\frac{u}{2}},$$

and $\lim_{N \rightarrow \infty} \varepsilon(N, f, r) = 0$.

Remark 5. It is part of that Theorem that $Z_N(f, r)$ is well-defined, at least if N is large enough (it remains unchanged under a modification of f on a zero Lebesgue measure set).

Remark 6. We shall prove a more precise version of the theorem, with explicit estimates on $\varepsilon(N, f, r)$; they will be useful to extend the validity of our results to probability densities which do not necessarily have finite moment of order 4, or finite L^p norm. Otherwise, it is sufficient to know that $\varepsilon(N, f, r) \rightarrow 0$ as $N \rightarrow \infty$.

The implications of Theorem 4 are best understood when recast in terms of the relative density of f with respect to γ ; so let

$$Z'_N(f, r) := \int_{S^{N-1}(r)} \left(\frac{f}{\gamma} \right)^{\otimes N} d\sigma_r^N. \quad (16)$$

Then, as a consequence of Theorem 4,

$$Z'_N(f, \sqrt{N}) = \frac{\sqrt{2}}{\Sigma} (1 + o(1)).$$

- *In other words, the integral of $f^{\otimes N}$ on $S^{N-1}(\sqrt{N})$ has a universal behavior – depending on f only through Σ .*

Thus, the fourth moment condition in Theorem 4 is just what is required, in the way of moments, for the conditioning to work. What about the L^p condition?

This comes in as follows: As a function of r , $Z_N(f, r)$ can be expressed in terms of the density for $\sum_{j=1}^N V_j^2$, where $\{V_j\}_{j \in \mathbb{N}}$ is a sequence of independent random variables with law $f(v)dv$; see Lemma 13 below. By Young's convolution inequality, the N -fold convolution power of a probability density g is continuous if g lies in $L^{N/(N-1)}$. Hence the L^p condition in Theorem 4 is natural: It is a simple sufficient condition to ensure that $Z_N(f, r)$ is a continuous function of r if N is large enough. Interestingly enough, though p can be arbitrarily close to 1, a bound on the entropy is *not* enough to ensure this. This point is discussed further in the appendix where we prove the version of the local central limit theorem that we shall use here.

When the conditions of Theorem 4 are satisfied, we may condition the tensor product $\mu^{\otimes N}$, with $\mu = f dv$, to obtain a probability measure on $S^{N-1}(\sqrt{N})$:

Definition 7 (Conditioned product measures). Given a probability density f on \mathbb{R} satisfying the hypotheses of Theorem 4, and $\mu(dv) = f(v)dv$, we define the corresponding family of *conditioned product measures*, denoted $\{[\mu^{\otimes N}]_{S^{N-1}(\sqrt{N})}\}_{N \in \mathbb{N}}$, by

$$[\mu^{\otimes N}]_{S^{N-1}(\sqrt{N})} := \frac{\prod_{j=1}^N f(v_j)}{Z_N(f, \sqrt{N})} \sigma^N = \frac{\prod_{j=1}^N (f(v_j)/\gamma(v_j))}{Z'_N(f, \sqrt{N})} \sigma^N.$$

The point of this definition is that, as noted above, one might hope that the family $\{[\mu^{\otimes N}]_{S^{N-1}(\sqrt{N})}\}_{N \in \mathbb{N}}$ would be μ -chaotic. This is indeed the case, and in a very strong sense, as we shall explain in the next subsection.

1.6. Entropic chaos. The notion of chaos as originally defined by Kac is well adapted to his original purpose, namely, establishing a rigorous connection between the linear Kac master equation on the one hand, and the non linear Boltzmann–Kac equation on the other. However, it is not quite strong enough to draw conclusions about the entropic rate of convergence to equilibrium for the Boltzmann–Kac equation from an analysis of the entropic rate of convergence for the Kac master equation. As we have explained above, a rigorous deduction in this direction would depend on having a precise version of the extensivity property (11) for chaotic families. Thus we ask:

Question 2. Is there a reformulation of the chaos property in entropic terms that is sufficiently strong that it can yield a bound on the entropic rate of convergence to equilibrium for (8) when combined with a bound on the entropic rate of convergence for (1)?

As we shall see, the answer is positive:

Definition 8 (Entropic μ -chaos). Let μ be a probability measure on \mathbb{R} , and, for each positive integer N , let $\mu^{(N)}$ be a probability measure on $S^{N-1}(\sqrt{N})$. The sequence $\{\mu^{(N)}\}_{N \in \mathbb{N}}$ is said to be entropically μ -chaotic in case it satisfies conditions (i) – (ii) in Definition 1, and in addition

$$(iii) \quad \lim_{N \rightarrow \infty} \frac{H(\mu^{(N)} | \sigma^N)}{N} = H(\mu | \gamma).$$

As indicated above, from the physical point of view, condition (iii) can be thought of as expressing *asymptotic extensivity* of the entropy for an entropically chaotic family; it provides a bridge between the entropy of the N -particle system and the entropy of the reduced system. This is reminiscent of a work by Kosygina on the limit from microscopic to macroscopic entropy in the Ginzburg-Landau model [19].

Secondly, entropic chaos really is a stronger notion than plain chaos; it involves *all* of variables, not only finite-dimensional marginals of fixed size. There is a good analogy with a work by Ben Arous and Zeitouni [3] (also based on the extensivity properties of entropy). Their work, just as ours, uses a version of the Central Limit Theorem.

Finally, once Condition (ii) is enforced, Condition (iii) really means that $\mu^{(N)}$ is “strongly” close to $\mu^{\otimes N}$. To understand this, think of the following well-known theorem: If $f^{(N)}$ is a family of probability densities on \mathbb{R} , converging weakly to some probability density f as $N \rightarrow \infty$, and J is a strictly convex functional, then automatically $J(f) \leq \liminf J(f^{(N)})$; but if in addition $J(f^{(N)}) \rightarrow J(f)$, then the convergence of $f^{(N)}$ to f actually holds almost in the sense of L^1 norm, not just weakly. So one could define a notion of strong convergence by requiring the weak convergence of $f^{(N)}$, plus the convergence of $J(f^{(N)})$ to $J(f)$. Such a step has already been taken in the definition of the “entropic convergence” used in the context of (deterministic) hydrodynamic limits of the Boltzmann equation by Golse and collaborators, in an impressive series of papers, starting with [1] and leading up to [16].

The following theorems provides an answer to both Questions 1 and 2:

Theorem 9. Let f be a probability density on \mathbb{R} satisfying

$$\int f(v)v^2 dv = 1 \quad \int f(v)v^4 dv < +\infty, \quad f \in L^\infty(\mathbb{R}),$$

and let $\mu(dv) = f(v)dv$. Then $\{[\mu^{\otimes N}]_{S^{N-1}(\sqrt{N})}\}_{N \in \mathbb{N}}$ is entropically μ -chaotic. In fact, condition (ii) from the definition of chaos holds in the following much stronger sense:

$$\lim_{N \rightarrow \infty} H(P_k([\mu^{\otimes N}]_{S^{N-1}(\sqrt{N})}) | \mu^{\otimes k}) = 0. \quad (17)$$

Furthermore, let $\{\mu^{(N)}\}_{N \in \mathbb{N}}$ be any family of symmetric probability measures on $S^{N-1}(\sqrt{N})$ such that

$$(H) \quad \frac{1}{N} H(\mu^{(N)} | [\mu^{\otimes N}]_{S^{N-1}(\sqrt{N})}) \xrightarrow{N \rightarrow \infty} 0. \text{ Then } \{\mu^{(N)}\}_{N \in \mathbb{N}} \text{ is entropically } \mu\text{-chaotic.}$$

Theorem 9 takes care of Question 1 for bounded densities f with a finite fourth moment, but certainly we cannot directly employ the conditioned tensor product construction when f does not have a fourth moment. However, using a diagonal

argument, we shall be able to show that there does exist an entropically $f dv$ -chaotic family for all finite energy, finite entropy probability densities f on \mathbb{R} :

Theorem 10. Let f be a probability density on \mathbb{R} with

$$\int f(v)v^2 dv = 1, \quad H(f|\gamma) < +\infty.$$

Then there exists an $f(v)dv$ -entropically chaotic sequence.

Theorem 9 has the following shortcoming: One might hope that for *any* $f dv$ -entropically chaotic family $\{\mu^{(N)}\}$, and not only conditioned tensor products, one would have

$$(ii') \text{ For any } k \in \mathbb{N}, \lim_{N \rightarrow \infty} H(P_k(\mu^{(N)}|\mu^{\otimes k})) = 0.$$

We would then include condition (ii') in the definition of entropic chaos. However, Theorem 9 asserts this only when $\{\mu^{(N)}\}$ is a conditioned tensor product. While the set of conditioned tensor product states is not propagated into itself by the Kac master equation, probability densities satisfying condition (H) for some $\{[\mu^{\otimes N}]_{S^{N-1}(\sqrt{N})}\}_{N \in \mathbb{N}}$ may well be. This leads to the following problem, for which we have no solution:

Open Problem 11. Does Condition (H) in Theorem 9 also imply Condition (ii')? More generally, does (ii') hold for a larger and easily recognized class of chaotic sequences, larger than those constructed by means of conditioning tensor products?

Also, as indicated above, a natural next step in the development of Kac's program consists in studying the propagation of Conditions (iii) or (H) — or (ii') — under the Kac master equation.

1.7. Final remarks. As an intermediate step in the proof of Theorem 9, we shall establish the following statement:

Theorem 12 (asymptotic upper semi-extensivity of the entropy). For each N , let $\mu^{(N)}$ be a probability density on $S^{N-1}(\sqrt{N})$, such that $\mu^{(N)}$ is μ -chaotic, in the sense of Definition 1. Then

$$H(\mu|\gamma) \leq \liminf_{N \rightarrow \infty} \frac{H(\mu^{(N)}|\sigma^N)}{N}.$$

This result certainly has interest in its own right, and further explains the meaning of Condition (H) in Theorem 9. By the way, Theorem 12 and Theorem 10 together provide a proof of Remark 2 following Theorem 6.1 in [26]. (The author had at the time thought that this remark was obvious.) This combination of results also establishes a kind of Γ -convergence of the functionals $H(\cdot|\sigma^N)/N$ to the functional $H(\cdot|\gamma)$.

We close our introductory discussion with some final remarks on Kac's program. Kac suggested that one could prove quantitative theorems on the non linear Boltzmann-Kac equation by means of an investigation of the linear master equation. At the time Kac wrote his paper, the rigorous mathematical theory of the Boltzmann equation had been in the doldrums since the landmark work of Carleman [6] in the thirties. The suggestion of Kac to recast the problem of investigating nonlinear equations such as (8) from a *probabilistic* many particle point of view was made in the hope that this might be a better path to progress.

However, the history of the subject has not developed as Kac had hoped. The lack of progress between the papers of Carleman and Kac turned out to be due as much to lack of attention as to the intrinsic difficulties of nonlinear equations such as (8). Once a new generation of mathematicians took up such equations as an active field of research, a well developed and full-fledged theory emerged. And so far, no relevant property of the nonlinear equation (8) has been proved via (1), which cannot be proved by direct means. Indeed, once again in this paper, *we shall prove lack of a uniform entropy production inequality for the Kac master equation (Theorem 3) through an analysis of the Boltzmann–Kac equation.*

Still, Kac’s program is worth trying to push for various reasons. First, the theory of spatially homogeneous Boltzmann equations has reached maturity, with quite precise results, and specialists are now looking for very sharp statements; it might be that Kac’s approach, thanks to its strong physical content, could be adapted to such refinements. Just because so far no relevant property of (8) has been first proved via (1) does not mean that this cannot be done, and certainly the probabilistic ground is less worked-over.

Second, it can be seen as a baby model for the much more subtle problem of propagation of chaos in the “true” spatially inhomogeneous Boltzmann equation.

Finally, one might be interested in it for just historical reasons, since Kac’s paper is one of the founding works in modern kinetic theory — and just perhaps, the renewed focus on Kac’s ideas will yield new progress of a fundamental sort.

1.8. Organization of the paper. In Section 2 below, we first study the asymptotics of the restricted tensor product, and prove Theorem 4. In Section 3, we establish the asymptotic upper semi-extensivity of the entropy (Theorem 12). In Section 4, we study the convergence of marginals, establishing in particular Condition (ii) of Definition 1 for the restricted tensor product. Asymptotic extensivity of the restricted tensor product (or perturbations thereof) will be proven in Section 5. Then, in Section 6 we prove Theorem 10. Finally, we shall investigate entropy production and prove Theorem 3 in Section 7. The appendix contains the statement and proof of a version of the local central limit theorem with precise quantitative bounds that we require in Section 2, but it also has some independent interest.

We close this introduction by thanking Julien Michel for providing reference [15]; and Alessio Figalli for his careful reading of and comments on an earlier version of the manuscript.

2. Asymptotics of the restricted tensor product. The goal of this section is to analyze the asymptotic behavior of

$$Z_N(f, r) := \int_{S^{N-1}(r)} f^{\otimes N} d\sigma_r^N$$

as $N \rightarrow \infty$, where σ_r^N is the uniform probability measure on $S^{N-1}(r)$.

Lemma 13 (probabilistic interpretation of Z_N). Let f be a probability density on \mathbb{R} , and let $\{V_j\}_{j \in \mathbb{N}}$ be a sequence of independent random variables with common law $f(v)dv$. Then the random variable $S_N := \sum_{j=1}^N V_j^2$ has density $s_N(u) du$, where

$$s_N(u) = \frac{|S^{N-1}|}{2} u^{\frac{N}{2}-1} Z_N(f, \sqrt{u}), \tag{18}$$

where $|S^{N-1}| = \frac{2\pi^{N/2}}{\Gamma(\frac{N}{2})}$ stands for the $(N-1)$ -dimensional volume of the unit sphere $S^{N-1} \subset \mathbb{R}^N$. As a particular case, the law of V_1^2 has density $h(u) du$, where

$$h(u) = \frac{1}{2\sqrt{u}} \left(f(\sqrt{u}) + f(-\sqrt{u}) \right). \quad (19)$$

Proof of Lemma 13. We use the notation $r = \sqrt{\sum v_i^2}$ and let \mathbb{E} denote the expectation with respect to the uniform probability measure on S^{N-1} . Whenever φ is a continuous test function supported in $[0, +\infty)$, a polar change of variables leads to

$$\begin{aligned} \mathbb{E} \varphi \left(\sum_{j=1}^N V_j^2 \right) &= \int_{\mathbb{R}^n} f^{\otimes N}(v) \varphi(r^2) dv \\ &= |S^{N-1}| \int_{[0, +\infty) \times S^{N-1}} f^{\otimes N} \varphi(r^2) r^{N-1} dr d\sigma \\ &= |S^{N-1}| \int_0^{+\infty} \varphi(u) \left(\frac{u^{\frac{N-1}{2}}}{2\sqrt{u}} \int_{S^{N-1}} f(\sqrt{u} y_1) \dots f(\sqrt{u} y_N) d\sigma(y) \right) du \\ &= \int_0^{+\infty} \varphi(u) \left(\frac{|S^{N-1}|}{2} u^{\frac{N}{2}-1} Z_N(f, \sqrt{u}) \right) du. \end{aligned}$$

□

Our next theorem, which is the main result of this section, is a slightly sharpened version of Theorem 4. It provides more information on how the remainder terms there depend on f . In describing this dependence, we shall use the following notation:

Let f be a probability density on \mathbb{R} with finite moment of order 4, and finite L^p norm for some $p \in (1, \infty)$. Define the mean kinetic energy and its variance by

$$E = \int f(v) v^2 dv; \quad \Sigma = \sqrt{\int_{\mathbb{R}} (v^2 - E)^2 f(v) dv}.$$

(As in the introduction, we have chosen units in which the mass m is equal to 2.) Let $\underline{E}, \bar{E}, \underline{\Sigma}, \bar{\Sigma}$ be constants such that

$$0 < \underline{E} \leq E \leq \bar{E} < +\infty; \quad \Sigma \geq \underline{\Sigma} > 0; \quad \|f\|_{L^p} \leq \bar{\Sigma},$$

and let χ_4 be any nonnegative function of $r > 0$, such that $\chi_4(r) \rightarrow 0$ as $r \rightarrow 0$ and

$$\int_{|v| \geq \frac{1}{r}} f(v) v^4 dv \leq \chi_4(r).$$

(For instance, we could define χ_4 to be the left hand side, but in applications another choice, such as $r^2 \int_{\mathbb{R}} f(v) v^6 dv$ may be more useful if, say, f possesses a 6th moment.)

Theorem 14 (asymptotics for the conditioned tensor product). With the above notation, define

$$Z_N(f, r) := \int_{S^{N-1}(r)} f^{\otimes N} d\sigma_r^N,$$

where $S^{N-1}(r)$ is the sphere of radius r in \mathbb{R}^N , and σ_r^N is the uniform probability measure on that sphere. Then, as $N \rightarrow \infty$,

$$Z_N(f, r) = \frac{\sqrt{2}}{\Sigma} \gamma^{(N)}(r) \left(\frac{\alpha_N(N)}{\alpha_N(r^2)} \right) \left[e^{-\frac{(r^2 - NE)^2}{2N\Sigma^2}} + o(1) \right],$$

where

$$\gamma^{(N)}(r) = \frac{e^{-r^2/2}}{(2\pi)^{N/2}}$$

is the restriction of $\gamma^{\otimes N}$ to $S^{N-1}(r)$,

$$\alpha_N(s) = s^{\frac{N}{2}-1} e^{-\frac{s}{2}},$$

and $o(1)$ stands for an expression which is bounded by a function $\omega(N) \rightarrow 0$, depending only on $\underline{E}, \overline{E}, \underline{\Sigma}, p, \overline{L}$ and χ_4 .

In particular,

$$Z_N(f, \sqrt{N}) = \frac{\sqrt{2}}{\Sigma} \gamma^{(N)}(\sqrt{N}) \left(e^{-\frac{N(1-E)^2}{2\Sigma^2}} + o(1) \right).$$

Proof of Theorem 14. Since (with the notation of Lemma 13) V_1^2 has density h , it follows that S_N has density h^{*N} (the N -fold convolution product of h with itself). So $s_N = h^{*N}$, which leads to the formula

$$Z_N(f, \sqrt{u}) = \frac{2h^{*N}(u)}{u^{\frac{N}{2}-1} |S^{N-1}|}. \quad (20)$$

We shall use the local central limit theorem to approximate h^{*N} . For that we need some estimates on h . First note that $\int h(u) u \, du = E$ and $\int h(u) u^2 \, du = E^2 + \Sigma^2$. Also,

$$\int_{u \geq 1/r} h(u) u^2 \, du = \int_{|v| \geq 1/\sqrt{r}} f(v) v^4 \, dv.$$

Next, let $q > 1$; by convexity of $t \mapsto t^q$, and the definition of h ,

$$\begin{aligned} \int_{\mathbb{R}_+} h^q(u) \, du &\leq \frac{1}{2} \int_{\mathbb{R}_+} u^{-q/2} (f^q(\sqrt{u}) + f^q(-\sqrt{u})) \, du \\ &= \int_{\mathbb{R}_+} u^{-(q-1)/2} \frac{1}{2\sqrt{u}} (f^q(\sqrt{u}) + f^q(-\sqrt{u})) \, du \\ &= \int_{\mathbb{R}} |v|^{1-q} f^q(v) \, dv \leq \int_{[-1,1]} \frac{f^q(v)}{|v|^{q-1}} \, dv + \int f^q(v) \, dv. \end{aligned}$$

If $q < (2p)/(p+1)$, then, by Hölder's inequality,

$$\int_{[-1,1]} \frac{f^q(v)}{|v|^{q-1}} \, dv \leq \left(\int f^p \right)^{\frac{q}{p}} \left(\int_{[-1,1]} \frac{dv}{|v|^{\frac{(q-1)p}{p-q}}} \right)^{1-\frac{q}{p}} \leq C(p, q) \|f\|_{L^p}^q.$$

On the other hand, $\int f^q \, dv = \|f\|_{L^q}^q \leq \|f\|_{L^p}^q$ as soon as $q \leq p$, because f is a probability measure. The conclusion is that there is a finite constant $C(p, q)$ such that

$$\|h\|_{L^q} \leq C(p, q) \|f\|_{L^p} \quad \text{for all } 1 < q < (2p)/(p+1). \quad (21)$$

Now let g be defined by $g(v) = \Sigma h(E + \Sigma v)$ so that

$$\int g(v) \, dv = 1 \quad \int g(v) v \, dv = 0 \quad \int g(v) v^2 \, dv = 1.$$

It follows immediately from (21) that g lies in L^q for some $q > 1$.

Also, g inherits from h a sort of “concentration bound” that we require to apply Theorem 27 in Appendix A:

$$\begin{aligned} \int_{u \geq \frac{1}{r}} g(u) u^2 du &= \frac{1}{\Sigma^2} \int_{s \geq \Sigma/r + E} h(s) (s - E)^2 ds \\ &\leq \frac{2}{\Sigma^2} \int_{s \geq \Sigma/r + E} h(s) (s^2 + E^2) ds \\ &\leq \frac{2}{\Sigma^2} \int_{s \geq \Sigma/r + E} h(s) s^2 ds + \frac{2}{\Sigma^2} \frac{E^2 + \Sigma^2}{(\Sigma/r + E)^2}. \end{aligned}$$

Evidently, the quantity on the left goes to 0 as $r \rightarrow 0$, with a rate which depends on $\int_{|v| \geq 1/\sqrt{r}} f(x) x^4 dx$.

As a conclusion, g satisfies all the assumptions of Theorem 27 in Appendix A, so there is a function $\lambda(N)$, only depending on the above-mentioned bounds, such that

$$\sup_{u \in \mathbb{R}} |\sqrt{N} g^{*N}(\sqrt{N}u) - \gamma(u)| \leq \lambda(N),$$

and so,

$$\sup_{u \in \mathbb{R}} \left| g^{*N}(u) - \frac{1}{\sqrt{N}} \gamma\left(\frac{u}{\sqrt{N}}\right) \right| \leq \frac{\lambda(N)}{\sqrt{N}}.$$

Then, since $g^{*N}(u) = \Sigma h^{*N}(NE + \Sigma u)$, we deduce

$$\sup_{x \in \mathbb{R}} \left| h^{*N}(x) - \frac{1}{\sqrt{N}\Sigma} \gamma\left(\frac{x - NE}{\sqrt{N}\Sigma}\right) \right| \leq \frac{\lambda(N)}{\sqrt{N}\Sigma}.$$

Now let us insert this bound in (20) and apply Stirling’s formula, in the form

$$\Gamma\left(\frac{N}{2}\right) = \sqrt{\pi N} \alpha_N(N) 2^{-\frac{N}{2}+1} \left(1 + O\left(\frac{1}{\sqrt{N}}\right)\right),$$

where $\alpha_N(u) := u^{\frac{N}{2}-1} e^{-u/2}$. This results in

$$Z_N(f, \sqrt{u}) = \frac{\sqrt{N} \alpha_N(N) 2^{-\frac{N}{2}+1}}{(\pi u)^{\frac{N}{2}-1} \sqrt{N}\Sigma} \left(\gamma\left(\frac{u - NE}{\sqrt{N}\Sigma}\right) + o(1) \right) \left(1 + O\left(\frac{1}{\sqrt{N}}\right)\right).$$

Now the desired expression follows easily. \square

Remark 15. Consider the case when $E = 1$; then

$$Z_N(f, \sqrt{N}) \sim \gamma^{(N)}(\sqrt{N}) \frac{\sqrt{2}}{\Sigma}.$$

Thus, after renormalization by the Gaussian density, $Z_N(f, \sqrt{N})$ has a nontrivial finite limit as $N \rightarrow \infty$. This was derived in the above proof by using successively the Local Central Limit Theorem and Stirling’s formula. Actually, the latter can be eliminated: Since $\gamma^{\otimes N}$ is constant on the sphere (by the way, the Gaussian is the only tensor product function to satisfy this property), we can write, with the notation (16),

$$Z'_N(f, r) = \frac{\int_{S^{N-1}(r)} f^{\otimes N} d\sigma_r^N}{\int_{S^{N-1}(r)} \gamma^{\otimes N} d\sigma_r^N} = \frac{Z_N(f, r)}{Z_N(\gamma, r)},$$

and in view of Lemma 13 this simplifies into

$$Z'_N(f, r) = \frac{h^{*N}(f, r^2)}{h^{*N}(\gamma, r^2)},$$

the asymptotics of which can be computed by using only the Local Central Limit Theorem, and not Stirling's formula. Actually, this line of reasoning shows how to *deduce* Stirling's formula from the Local Central Limit Theorem. This observation in itself is not new (it is made explicitly in the discussion of the Local Central Limit theorem in [15]; see also [21, Problem 2]). However, it is interesting to see that here it arises naturally as part of a physically relevant problem.

We conclude this section with a higher-dimensional generalization of Theorem 14, beginning with the analogs of some definitions we made for densities on \mathbb{R} . Let f be a probability density on \mathbb{R}^k with finite moment of order 4, and finite L^p norm for some $p \in (1, \infty)$. Define

$$E = \int_{\mathbb{R}^k} f(v)|v|^2 dv; \quad \Sigma = \sqrt{\int_{\mathbb{R}^k} (|v|^2 - E)^2 f(v) dv}.$$

Let $\underline{E}, \overline{E}, \underline{\Sigma}, \overline{\Sigma}$ be such that

$$0 < \underline{E} \leq E \leq \overline{E} < +\infty; \quad \Sigma \geq \underline{\Sigma} > 0; \quad \|f\|_{L^p} \leq \overline{\Sigma},$$

and let χ_4 be a nonnegative function of $r > 0$, such that $\chi_4(r) \rightarrow 0$ as $r \rightarrow 0$ and

$$\int_{|v| \geq \frac{1}{r}} f(v)|v|^4 dv \leq \chi_4(r).$$

Again in the higher dimensional case, it is easy to see that if $f \in L^p$ for some $p > 1$, then the density h of v^2 under $f(v)dv$ is in L^q for some $q > 1$. Then we have the following analogue of Theorem 14:

Theorem 16 (asymptotics for the conditioned tensor product). With the above notation, define

$$Z_m(f, r) := \int_{S^{km-1}(r)} f^{\otimes m} d\sigma_r^{km},$$

Then, as $m \rightarrow \infty$,

$$Z_m(f, r) = \frac{\sqrt{2}}{\Sigma} \frac{e^{-\frac{r^2}{2}}}{(2\pi)^{\frac{km}{2}}} \left(\frac{\alpha_{km}(km)}{\alpha_{km}(r^2)} \right) \left[e^{-\frac{(r^2 - mE)^2}{2m\Sigma^2}} + o(1) \right],$$

where $o(1)$ stands for an expression which is bounded by a function $\omega(m) \rightarrow 0$, depending only on $k, \underline{E}, \overline{E}, \underline{\Sigma}, p, \overline{\Sigma}$ and χ_4 .

In particular,

$$Z_m(f, r) = \frac{\sqrt{2}}{\Sigma} \gamma^{(km)}(r) \left(\frac{\alpha_{km}(km)}{\alpha_{km}(r^2)} \right) \left(e^{-\frac{(r^2 - mE)^2}{2\Sigma^2}} + o(1) \right).$$

Sketch of a proof of Theorem 16. First, one can adapt the proof of Lemma 13 to the present case; the conclusion should be changed as follows: $S_m := \sum_{j=1}^m |V_j|^2$ has density $s_m(u) du$, where

$$s_m(u) = \frac{|S^{km-1}|}{2} u^{\frac{km}{2}-1} Z_m(f, \sqrt{u}).$$

In particular, the law of $|V_1|^2$ is

$$h(u) = \frac{|S^{km-1}|}{2} u^{\frac{km}{2}-1} \int_{S^{k-1}(\sqrt{u})} f d\sigma_{\sqrt{u}}^k.$$

Then the proof of Theorem 14 adapts to the present case with hardly any change, upon replacement of N by mk . \square

3. Asymptotic upper semi-extensivity of the entropy. To motivate this section, let us recall an important property of the entropy functional. Consider a sequence of probability measures μ^N on \mathbb{R} , converging weakly to another probability distribution $\mu(dv)$ as $N \rightarrow \infty$, and let ν be another probability measure. Then

$$H(\mu|\nu) \leq \liminf_{N \rightarrow \infty} H(\mu^N|\nu).$$

In other words, the relative entropy is *lower semi-continuous* under weak convergence.

In this section, we shall show that the same property is true when the dimension goes to ∞ , and weak convergence is replaced by the chaos property. The following theorem is a generalization of Theorem 12.

Theorem 17. Let g be a probability density on \mathbb{R} , such that

$$\int g(x)x^2 dx = 1, \quad \int g(x)x^4 dx < +\infty, \quad g \in L^p(\mathbb{R}) \quad (p > 1);$$

define $\nu(dv) = g(v) dv$. For each positive integer N , let $\mu^{(N)}$ be a symmetric probability measure on $S^{N-1}(\sqrt{N})$, such that

$$P_1 \mu^{(N)} \xrightarrow{N \rightarrow \infty} \mu,$$

in the sense of weak convergence against bounded continuous functions. Then

$$H(\mu|\nu) \leq \liminf_{N \rightarrow \infty} \frac{H(\mu^{(N)} | [\nu^{\otimes N}]_{S^{N-1}(\sqrt{N})})}{N}.$$

More generally, if, for some positive integer k ,

$$P_k \mu^{(N)} \xrightarrow{N \rightarrow \infty} \mu_k,$$

then

$$\frac{H(\mu_k | \nu^{\otimes k})}{k} \leq \liminf_{N \rightarrow \infty} \frac{H(\mu^{(N)} | [\nu^{\otimes N}]_{S^{N-1}(\sqrt{N})})}{N}.$$

The proof uses the results of Section 2, plus a duality formula for the entropy:

Lemma 18 (Legendre representation of the H functional). Let \mathcal{X} be a locally compact complete metric space equipped with a reference (Borel) probability measure ν . Then, for any other probability measure μ on \mathcal{X} ,

$$H(\mu|\nu) = \sup \left\{ \int \varphi d\mu - \log \left(\int e^\varphi d\nu \right); \varphi \in C_b(\mathcal{X}) \right\}, \quad (22)$$

where $C_b(\mathcal{X})$ stands for the space of bounded continuous functions on \mathcal{X} .

Moreover, one can restrict the supremum in (22) to those functions φ such that $\int e^\varphi d\nu = 1$.

We skip the proof of Formula (22), which belongs to folklore (see e.g. [13], or [20, Appendix B] for a complete proof in the case of a compact space \mathcal{X}). As for the last part of Lemma 18, it follows from an easy homogeneity argument.

Proof of Theorem 17. Let $\nu^{(N)} := [\nu^{\otimes N}]_{S^{N-1}(\sqrt{N})}$. We first consider the case $k = 1$. Let $\varepsilon > 0$ be given. By Lemma 18, we can find a bounded continuous function φ such that

$$\int e^\varphi g = 1; \quad \int \varphi d\mu \geq H(\mu|\nu) - \varepsilon.$$

On $S^{N-1}(\sqrt{N})$, consider the function

$$\Phi(v_1, \dots, v_N) = \varphi(v_1) + \dots + \varphi(v_N).$$

By Lemma 18 again,

$$\frac{H(\mu^{(N)}|\nu^{(N)})}{N} \geq \frac{1}{N} \int_{S^{N-1}(\sqrt{N})} \Phi(v) d\mu^{(N)}(v) - \frac{1}{N} \log \left(\int e^{\Phi(v)} d\nu^{(N)}(v) \right). \quad (23)$$

The first term in the right-hand side of (23) is controlled by symmetry and convergence of the first marginal:

$$\frac{1}{N} \int_{S^{N-1}(\sqrt{N})} \Phi(v) d\mu^{(N)}(v) = \int_{\mathbb{R}} \varphi(v) d(P_1 \mu^{(N)})(v) \xrightarrow{N \rightarrow \infty} \int_{\mathbb{R}} \varphi d\mu. \quad (24)$$

(Here we have used the continuity of φ .)

To estimate the second term in the right-hand side of (23), we note that

$$\int e^{\Phi(v)} d\nu^{(N)}(v) = \frac{Z_N(e^\varphi g, \sqrt{N})}{Z_N(g, \sqrt{N})} = \frac{Z'_N(e^\varphi g, \sqrt{N})}{Z'_N(g, \sqrt{N})}. \quad (25)$$

Since φ is bounded above, we know that $e^\varphi g$ satisfies the same estimates as g , which makes it possible to apply Theorem 14: with obvious notation,

$$Z'_N(e^\varphi g, \sqrt{N}) = \frac{\sqrt{2}}{\Sigma(e^\varphi g)} \left(e^{-\frac{N(1-E(e^\varphi g))^2}{2\Sigma(e^\varphi g)^2}} + o(1) \right) = O(1).$$

Hence

$$\liminf_{N \rightarrow \infty} \left(-\frac{1}{N} \log Z'_N(e^\varphi g, \sqrt{N}) \right) \geq 0. \quad (26)$$

Similarly,

$$Z'_N(g, \sqrt{N}) = \frac{\sqrt{2}}{\Sigma(g)} \left(e^{-\frac{N(1-E(g))^2}{2\Sigma(g)^2}} + o(1) \right) = \frac{\sqrt{2}}{\Sigma(g)} (1 + o(1)),$$

so

$$\lim_{N \rightarrow \infty} \left(\frac{1}{N} \log Z'_N(g, \sqrt{N}) \right) = 0. \quad (27)$$

(This is where the assumption $\int g(x)x^2 dx = 1$ is used.)

The combination of (25), (26) and (27) implies

$$\liminf_{N \rightarrow \infty} \left(-\frac{1}{N} \log \left(\int e^{\Phi(v)} d\nu^{(N)}(v) \right) \right) \geq 0.$$

Combining this with (24) and (23), we find

$$\liminf_{N \rightarrow \infty} \frac{H(\mu^{(N)}|\nu^{(N)})}{N} \geq \int_{\mathbb{R}} \varphi d\mu,$$

and by the choice of φ this is no less than $H(\mu|\nu) - \varepsilon$. Since ε is arbitrarily small, the proof of Theorem 17 is complete in the case $k = 1$.

Now the proof for general k goes along the same lines: pick up $\varphi = \varphi(v_1, \dots, v_k)$ such that

$$\int_{\mathbb{R}^k} e^\varphi g^{\otimes k} = 1; \quad \int \varphi d\mu_k \geq H(\mu_k | \nu^{\otimes k}) - \varepsilon.$$

Define m as the integer part of N/k . On $S^{N-1}(\sqrt{N})$, consider the function

$$\Phi(v_1, \dots, v_N) = \varphi(v_1, \dots, v_k) + \varphi(v_{k+1}, \dots, v_{2k}) + \dots + \varphi(v_{(m-1)k+1}, \dots, v_{mk}).$$

Then (23) is unchanged, and (24) transforms into

$$\frac{1}{N} \int_{S^{N-1}(\sqrt{N})} \Phi(v) d\mu^{(N)}(v) = \left(\frac{m}{N}\right) \int_{\mathbb{R}^k} \varphi(v) d(P_k \mu^{(N)})(v) \xrightarrow{N \rightarrow \infty} \frac{1}{k} \int_{\mathbb{R}^k} \varphi d\mu_k. \quad (28)$$

Also equation (27) is unchanged; but there is a subtlety with equation (26), which cannot be directly interpreted in terms of constrained tensor product. So write $N = km + q$ ($0 \leq q \leq k-1$), and

$$x = (y, z), \quad y = (x_1, \dots, x_{km}), \quad z = (x_{km+1}, \dots, x_N).$$

By using polar changes of variables (successively for x and y) and a test function argument, we see that

$$\begin{aligned} & \int_{S^{N-1}(\sqrt{N})} F(y, z) d\sigma_r^N(y, z) \\ &= \int_{\mathbb{R}^q} \frac{r}{\sqrt{r^2 - |z|^2}} \left(\int_{S^{km-1}(\sqrt{r^2 - |z|^2})} F(y, z) \frac{|S^{km-1}(\sqrt{r^2 - |z|^2})|}{|S^{N-1}(r)|} d\sigma_{\sqrt{r^2 - |z|^2}}^{km} \right) dz, \end{aligned}$$

where the integral is restricted to the region $|z| \leq r$. Then we recognize that

$$\int_{S^{km-1}(\sqrt{r^2 - |z|^2})} e^\Phi(y) \left(\frac{g}{\gamma}\right)^{\otimes km}(y) d\sigma_{\sqrt{r^2 - |z|^2}}^{km-1}(y) = Z'_m(e^\varphi g^{\otimes m}, \sqrt{r^2 - |z|^2}).$$

So in the end

$$\begin{aligned} & \int_{S^{N-1}(\sqrt{N})} e^\Phi g^{\otimes N} d\sigma^N \\ &= \int_{\mathbb{R}^q} \sqrt{\frac{N}{N - |z|^2}} \frac{|S^{km-1}(\sqrt{N - |z|^2})|}{|S^{N-1}(\sqrt{N})|} \left(\frac{g}{\gamma}\right)^{\otimes q}(z) Z'_m(e^\varphi g, \sqrt{N - |z|^2}) dz, \end{aligned} \quad (29)$$

where the integral is restricted to the region $|z| \leq \sqrt{N}$. To conclude along the lines of the case $k = 1$, it is sufficient to show that the expression (29) is uniformly bounded as $N \rightarrow \infty$ (φ and g being fixed). Since there are a finite number of possible values for q , we might also assume that q is fixed.

Now use the formulas

$$\begin{aligned} Z'_m(e^\varphi g, \sqrt{N - |z|^2}) &= \left(\frac{\alpha_{km}(km)}{\alpha_{km}(N)}\right) \times O(1), \\ |S^{N-1}(r)| &= \frac{(2\pi)^{N/2} r^{N-1}}{\sqrt{\pi N} \alpha_N(N)} (1 + o(1)) \end{aligned}$$

to estimate (29): after some computations, one finds

$$\int_{S^{N-1}(\sqrt{N})} e^\Phi g^{\otimes N} d\sigma^N \leq \frac{O(1)}{(2\pi)^{q/2}} \int_{\mathbb{R}^q} \left(1 - \frac{|z|^2}{N}\right)_+^{\frac{km}{2} - 1} \left(\frac{g}{\gamma}\right)^{\otimes q} dz. \quad (30)$$

Next, if $|z| \leq \sqrt{N}$, then

$$\left(1 - \frac{|z|^2}{N}\right)^N \leq e^{-|z|^2},$$

so

$$\left(1 - \frac{|z|^2}{N}\right)_+^{\frac{km}{2}-1} \leq e^{-\frac{|z|^2}{2}\left(\frac{km-2}{N}\right)} 1_{|z| \leq \sqrt{N}} = e^{-\frac{|z|^2}{2}\left(1 - \frac{2+q}{N}\right)} 1_{|z| \leq \sqrt{N}}.$$

Plug this in (30) to obtain

$$\begin{aligned} \int_{S^{N-1}(\sqrt{N})} e^\Phi g^{\otimes N} d\sigma^N &\leq O(1) \int_{|z| \leq \sqrt{N}} e^{\frac{|z|^2}{N}\left(\frac{2+q}{N}\right)} g^{\otimes q}(z) dz \\ &\leq O(1) e^{1+\frac{q}{2}} \int_{\mathbb{R}^q} g^{\otimes q}(z) dz = O(1). \end{aligned}$$

With this estimate in hand, there is no difficulty to conclude the proof of Theorem 17. \square

4. Convergence of marginals. This section is devoted to the convergence of finite-dimensional marginals under various entropy assumptions. In the first subsection, we show that (with loose notation) the natural condition $H(\mu^{(N)}|\mu^{\otimes N}) = o(N)$ implies that $\mu^{(N)}$ is μ -chaotic. In the second subsection, we show that at least in the case when $\mu^{(N)}$ is the constrained tensor product, then the convergence of the finite-dimensional marginals holds in a stronger sense, namely in relative entropy (and as a consequence in total variation).

4.1. From entropy estimates to chaos.

Theorem 19 (Entropic closeness to the constrained tensor product implies chaos). Let $\nu(dv) = g(v) dv$ be a probability measure on \mathbb{R} , such that

$$\int g(v) v^2 dv = 1, \quad \int g(v) v^4 dv < +\infty, \quad g \in L^p(\mathbb{R}) \quad (p > 1),$$

and let $\nu^{(N)} = [\nu^{\otimes N}]_{S^{N-1}(\sqrt{N})}$ be the constrained tensor product of ν on $S^{N-1}(\sqrt{N})$. Let further $\mu^{(N)}$ be a symmetric probability measure on $S^{N-1}(\sqrt{N})$, such that

$$\frac{H(\mu^{(N)}|\nu^{(N)})}{N} \xrightarrow{N \rightarrow \infty} 0.$$

Then, $\mu^{(N)}$ is ν -chaotic. More precisely, for each k , the marginal $P_k \mu^{(N)}$ converges weakly (against bounded continuous test functions) to $\nu^{\otimes k}$.

Proof. We first claim that, for given k , the sequence $P_k \mu^{(N)}$ is tight. Indeed, let m be the integer part of N/k , then

$$\begin{aligned} N = \int |x|^2 d\mu^{(N)}(x) &\geq \int (x_1^2 + \dots + x_k^2) d\mu^{(N)}(x) + \dots + \\ &\quad \int (x_{(m-1)k+1}^2 + \dots + x_{mk}^2) d\mu^{(N)}(x), \quad (31) \end{aligned}$$

and by symmetry the latter expression is

$$m \int_{\mathbb{R}^k} |x|^2 d(P_k \mu^{(N)})(x).$$

It follows that

$$\int_{\mathbb{R}^k} |x|^2 d(P_k \mu^{(N)})(x) \leq \frac{N}{m},$$

which converges to k as $N \rightarrow \infty$.

By Prokhorov's theorem, $P_k \mu^{(N)}$ converges, possibly up to extraction of a subsequence, to some probability measure μ_k on \mathbb{R}^k . From Theorem 17,

$$\frac{H(\mu_k | \nu^{\otimes k})}{k} \leq \liminf_{N \rightarrow \infty} \frac{H(\mu^{(N)} | [\nu^{\otimes N}]_{S^{N-1}(\sqrt{N})})}{N} = 0.$$

It follows that $\mu_k = \nu^{\otimes k}$, so the whole sequence $P_k \mu^{(N)}$ does converge to $\nu^{\otimes k}$, and $(\mu^{(N)})$ is indeed ν -chaotic. \square

4.2. Marginals of the constrained tensor product. As an obvious consequence of Theorems 17 and 19, the constrained tensor product $\mu^{(N)}$ of μ is itself μ -chaotic. We shall show in this section a stronger result: $P_k \mu^{(N)}$ converges to $\mu^{\otimes k}$ in total variation, and even in relative entropy.

Theorem 20 (Property (ii') for the constrained tensor product). Let $\mu(dv) = f(v) dv$ be a probability measure on \mathbb{R} , such that $f \in L^p(\mathbb{R})$ for some $p > 1$, and $\int_{\mathbb{R}} v^4 f(v) dv < \infty$. Let $\mu^{(N)} = [\mu^{\otimes N}]_{S^{N-1}(\sqrt{N})}$ be the restricted N -fold tensor product of μ . Then for all positive integers k ,

$$\lim_{N \rightarrow \infty} H(P_k \mu^{(N)} | \mu^{\otimes k}) = 0.$$

Proof. Let $[f^{\otimes N}]_{S^{N-1}(\sqrt{N})}$ stand for the density of the constrained tensor product, with respect to the uniform probability measure σ^N . Fix any integer k . Then for all N sufficiently large,

$$[f^{\otimes N}]_{S^{N-1}(\sqrt{N})} = \left(\prod_{j=1}^k (f(v_j) / \gamma(v_j)) \right) \frac{\prod_{j=k+1}^N (f(v_j) / \gamma(v_j))}{Z'_N(f, \sqrt{N})}$$

With the notation $s^2 = \sum_{j=1}^k v_j^2$, this expression can be rewritten as

$$[f^{\otimes N}]_{S^{N-1}(\sqrt{N})} = \left(\prod_{j=1}^k (f(v_j) / \gamma(v_j)) \right) \frac{Z'_{N-k}(f, \sqrt{N-s^2})}{Z'_N(f, \sqrt{N})} \frac{\prod_{j=k+1}^N (f(v_j) / \gamma(v_j))}{Z'_{N-k}(f, \sqrt{N-s^2})}.$$

So

$$P_k(\mu^{(N)}) = \left(\prod_{j=1}^k (f(v_j) / \gamma(v_j)) \right) \frac{Z'_{N-k}(f, \sqrt{N-s^2})}{Z'_N(f, \sqrt{N})} P_k(\sigma^N). \quad (32)$$

As a consequence, with \mathcal{L}_k standing for the k -dimensional Lebesgue measure,

$$\begin{aligned} H(P_k \mu^{(N)} | \mu^{\otimes k}) &= \int_{\mathbb{R}^k} \left(\log \frac{dP_k(\mu^{(N)})}{d\mathcal{L}_k} - \log f^{\otimes k} \right) d(P_k \mu^{(N)}) \\ &= \int_{\mathbb{R}^k} \log \left(\frac{Z'_{N-k}(f, \sqrt{N-s^2})}{Z'_{N-k}(f, \sqrt{N})} \right) d(P_k \mu^{(N)}) \\ &\quad + \int_{\mathbb{R}^k} \log \frac{d(P_k \sigma^N)}{d\gamma^{\otimes k}} d(P_k \mu^{(N)}) \end{aligned}$$

From Theorem 14 and some computation,

$$\frac{Z'_{N-k}(f, \sqrt{N-s^2})}{Z'_{N-k}(f, \sqrt{N})} = \left(e^{-s^4 / (2\Sigma^2 N)} \right) (1 + o(1))$$

and so

$$H(P_k \mu^{(N)} | \mu^{\otimes k}) = \int_{\mathbb{R}^k} \log \frac{d(P_k \sigma^N)}{d\gamma^{\otimes k}} d(P_k \mu^{(N)}) + o(1).$$

In other words,

$$H(P_k \mu^{(N)} | \mu^{\otimes k}) = \int \Psi_N(y) f^{\otimes k}(y) dy + o(1), \quad (33)$$

where

$$\Psi_N(y) := \left(\frac{P_k(\sigma^N)}{\gamma^{\otimes k}} \right) \log \left(\frac{P_k(\sigma^N)}{\gamma^{\otimes k}} \right) \left(\frac{Z'_{N-k}(f, \sqrt{N - |y|^2})}{Z'_{N-k}(f, \sqrt{N})} \right). \quad (34)$$

Now let us derive some estimates on Ψ_N . By direct computation,

$$(P_k \sigma^N)(dy) = \frac{|S^{N-k-1}|}{N^{k/2} |S^{N-1}|} \left(1 - \frac{|y|^2}{N} \right)^{(N-k-2)/2} dy.$$

By an application of Stirling's formula, and some computation again,

$$\frac{d(P_k \sigma^N)}{d\gamma^{\otimes k}}(y) \leq (1 + o(1)) e^{\frac{(k+1)|y|^2}{N}} 1_{|y| \leq \sqrt{N}} \leq (1 + o(1)) e^{k+1} = O(1). \quad (35)$$

On the other hand, Theorem 14 implies that the ratio of the Z' terms in (34) is uniformly bounded. We conclude that $\Psi_N(y)$ itself is bounded above, uniformly in N and y . This makes it possible to apply the dominated convergence theorem, in the form

$$\limsup_{N \rightarrow \infty} \int \Psi_N(y) f^{\otimes k}(y) dy \leq \int \left(\limsup_{N \rightarrow \infty} \Psi_N(y) \right) f^{\otimes k}(y) dy.$$

It follows from (35) and Theorem 14 that for any $y \in \mathbb{R}^k$,

$$\begin{cases} \frac{dP_k \sigma^N}{d\gamma^{\otimes k}}(y) \longrightarrow 1 \\ \frac{Z'_{N-k}(f, \sqrt{N - |y|^2})}{Z'_{N-k}(f, \sqrt{N})} \longrightarrow 1 \end{cases}$$

as $N \rightarrow \infty$. So $\lim \Psi_N(y) = 0$, and as a consequence $\limsup_{N \rightarrow \infty} \int \Psi_N(y) f^{\otimes k}(y) dy \leq 0$, so by (33), $\limsup_{N \rightarrow \infty} H(P_k \mu^{(N)} | \mu^{\otimes k}) \leq 0$. This concludes the proof of Theorem 20. \square

5. From microscopic to macroscopic entropy. Now comes one of the main results of this paper.

Theorem 21. Let f be a probability density on \mathbb{R} , such that

$$\int f(v) v^2 dv = 1 \quad \int f(v) v^4 dv < +\infty \quad f \in L^\infty(\mathbb{R}).$$

Let $\nu(dv) = f(v) dv$, and let $\nu^{(N)} = [\nu^{\otimes N}]_{S^{N-1}(\sqrt{N})}$ be the constrained N -fold tensor product of ν . For each N , let further $\mu^{(N)}$ be a probability measure on $S^{N-1}(\sqrt{N})$ such that

$$\frac{H(\mu^{(N)} | \nu^{(N)})}{N} \xrightarrow{N \rightarrow \infty} 0.$$

Then

$$\frac{H(\mu^{(N)}|\sigma^N)}{N} \xrightarrow{N \rightarrow \infty} H(\nu|\gamma).$$

Remark 22. During the proof, we shall show that the convergence of the marginals $P_k \mu^{(N)}$ actually holds true in the sense of weak convergence against bounded measurable functions (as opposed to bounded continuous functions). We do not know whether it holds true in the sense of, say, total variation.

Proof of Theorem 21. First we write

$$\begin{aligned} H(\mu^{(N)}|\sigma^N) &= \int \log \frac{d\mu^{(N)}}{d\sigma^N} d\mu^{(N)} \\ &= \int \log \frac{d\mu^{(N)}}{d\nu^{(N)}} d\mu^{(N)} + \int \log \frac{d\nu^{(N)}}{d\sigma^N} d\mu^{(N)} \\ &= H(\mu^{(N)}|\nu^{(N)}) + \int \log \left(\frac{f}{\gamma} \right)^{\otimes N} d\mu^{(N)} - \log(Z'_N(f, \sqrt{N})) \\ &= o(N) + N \int \log f(v_1) d\mu^{(N)}(v) - \int \log \gamma^{\otimes N} d\mu^{(N)} - \log Z'_N(f, \sqrt{N}) \\ &= o(N) + N \int \log f(v_1) d\mu^{(N)}(v) + N \left(\frac{1 + \log(2\pi)}{2} \right). \end{aligned}$$

In the next to last step, we have used Theorem 14, which implies that $\log Z'_N(f, \sqrt{N})$ converges to a positive limit as $N \rightarrow \infty$, and so may be absorbed into the $o(N)$ term. Also, in the last step we have replaced $\gamma^{\otimes N}$ by its explicit expression on $S^{N-1}(\sqrt{N})$.

Then, after division by N , we find

$$\begin{aligned} \frac{H(\mu^{(N)}|\sigma^N)}{N} &= \int \log f(v_1) d\mu^{(N)}(v) + \left(\frac{1 + \log(2\pi)}{2} \right) + o(1) \\ &= \int_{\mathbb{R}} \log f(v_1) d(P_1 \mu^{(N)})(v_1) + \left(\frac{1 + \log(2\pi)}{2} \right) + o(1). \end{aligned}$$

For any $\delta > 0$, we have therefore

$$\frac{H(\mu^{(N)}|\sigma^N)}{N} \leq \int \log(f(v_1) + \delta) d(P_1 \mu^{(N)})(v_1) + \left(\frac{1 + \log(2\pi)}{2} \right) + o(1). \quad (36)$$

Assume for the moment that f is continuous. Then $\log(f + \delta)$ is a bounded continuous function, so we can pass to the limit, using the weak convergence of $P_1 \mu^{(N)}$ to $\nu(dv) = f(v) dv$ (Theorem 19), and deduce

$$\limsup_{N \rightarrow \infty} \frac{H(\mu^{(N)}|\sigma^N)}{N} \leq \int f(v_1) \log(f(v_1) + \delta) dv_1 + \left(\frac{1 + \log(2\pi)}{2} \right). \quad (37)$$

By dominated convergence, we can now let $\delta \rightarrow 0$, and recover

$$\limsup_{N \rightarrow \infty} \frac{H(\mu^{(N)}|\sigma^N)}{N} \leq \int f(v_1) \log f(v_1) dv_1 + \left(\frac{1 + \log(2\pi)}{2} \right).$$

Since $\int f(v)v^2 dv = 1$, it is easy to check that the latter expression coincides with $H(\nu|\gamma)$. The conclusion is that

$$\limsup_{N \rightarrow \infty} \frac{H(\mu^{(N)}|\sigma^N)}{N} \leq H(\nu|\gamma). \quad (38)$$

On the other hand, by Theorem 17, applied with $g = \gamma$,

$$H(\mu|\gamma) \leq \liminf_{N \rightarrow \infty} \frac{H(\mu^{(N)}|\sigma^N)}{N}. \quad (39)$$

The combination of (38) and (39) concludes the proof of Theorem 21. \square

Now, let us prove the more general statement alluded to in Remark 22. We start again from (36), and deduce that

$$\frac{H(\mu^{(N)}|\sigma^N)}{N} \leq \log \|f\|_{L^\infty} + \left(\frac{1 + \log(2\pi)}{2} \right) + o(1),$$

which is bounded as $N \rightarrow \infty$. This bound can be combined with the exact (not asymptotic) inequality

$$\frac{H(P_k \mu^{(N)}|P_k \sigma^N)}{k} \leq 2 \frac{H(\mu^{(N)}|\sigma^N)}{N}, \quad (40)$$

to obtain

$$H(P_k \mu^{(N)}|P_k \sigma^N) = O(1).$$

The inequality (40) is a generalization of the subadditivity inequality on S^N from [10], which gives the $k = 1$ case. The generalization to higher k can be found in [2], in Example 1 under Corollary 5 there.

Next, by the same kind of computation as in the beginning of the proof,

$$H(P_k \mu^{(N)}|\gamma^{\otimes k}) = H(P_k \mu^{(N)}|P_k \sigma^N) + \int \log \frac{d(P_k \sigma^N)}{d\gamma^{\otimes k}} d(P_k \mu^{(N)}).$$

It follows by (35) that

$$H(P_k \mu^{(N)}|\gamma^{\otimes k}) \leq H(P_k \mu^{(N)}|P_k \sigma^N) + C,$$

where C is some constant depending only on k . In particular, $H(P_k \mu^{(N)}|\gamma^{\otimes k})$ is bounded as $N \rightarrow \infty$. The conclusion is that *the marginals $P_k \mu^{(N)}$ have bounded relative entropy with respect to $\gamma^{\otimes k}$, uniformly in N* . It follows by the Dunford-Pettis compactness criterion that the densities $f_k^{(N)}$ of $P_k \mu^{(N)}$ constitute a compact set in $L^1(\mathbb{R}^k)$, equipped with the weak topology. Since this family converges weakly to $f^{\otimes k}$ as $N \rightarrow \infty$, actually the limit

$$\int_{\mathbb{R}^k} \psi(v) f_k^{(N)}(v) dv \xrightarrow{N \rightarrow \infty} \int_{\mathbb{R}^k} \psi(v) f^{\otimes k}(v) dv$$

holds true for all bounded measurable functions ψ , not necessarily continuous. The conclusion follows by the same arguments as before.

6. Generalization to unbounded densities. In this section we use a density argument to derive Theorem 10 from Theorem 21.

Proof of Theorem 10. If f is bounded and has a finite fourth moment, we can simply use the tensor product construction. Otherwise, we define approximations to f as follows: If f has a finite fourth moment but is unbounded, and $\delta > 0$, define f_δ to be $e^{\delta \Delta} f$, rescaled so that f_δ has unit variance. Otherwise, if f does not have a finite fourth moment, let

$$g_\delta = e^{\delta \Delta} (f 1_{[-1/\delta, 1/\delta]}).$$

Then renormalize g_δ so that it is a probability density, and finally, make an affine change of variable to obtain a density that has zero mean and unit variance. Call this f_δ .

It is easy to see that for any positive integer j , we can choose a value $\delta_j > 0$ so that

$$|H(f_{\delta_j}|\gamma) - H(f|\gamma)| < \frac{1}{2j}. \quad (41)$$

Apply the tensor product construction with each f_{δ_j} to produce the chaotic sequence $\mu_{\delta_j}^{(N)}$. By Theorem 21,

$$\frac{1}{N}H(\mu_{\delta_j}^{(N)}|\sigma^N) \xrightarrow{N \rightarrow \infty} H(f_{\delta_j}|\gamma).$$

Therefore, we may inductively define an increasing sequence of integers $\{N_j\}$ by choosing $N_j > N_{j-1}$ large enough that

$$\left| \frac{1}{N}H(\mu_{\delta_j}^{(N)}|\sigma^N) - H(f_{\delta_j}|\gamma) \right| < \frac{1}{2j} \quad (42)$$

for all $N > N_j$.

Combining (41) and (42), we obtain

$$\left| \frac{1}{N}H(\mu_{\delta_j}^{(N)}|\sigma^N) - H(f|\gamma) \right| < \frac{1}{j}. \quad (43)$$

Further increasing the N_j if required, we may assume, on account of Theorem 20, that for each j ,

$$N \geq N_j \implies \sup_{1 \leq \ell \leq j} H(P_\ell \mu_{\delta_j}^{(N)} | f_{\delta_j}^{\otimes \ell}) < \frac{1}{2j^2}. \quad (44)$$

We are now ready to define our sequence, which we shall show to be $f(v)dv$ -chaotic in the entropic sense: For each N , define

$$\mu^{(N)} = \mu_{\delta_k}^{(N_k)} \quad \text{for} \quad k = \sup\{\ell : N_\ell < N\}.$$

First, property (i) holds for obvious reasons. Next, to see that property (ii) holds, let ϕ be any continuous bounded function on \mathbb{R}^k . Then, by the well-known Csiszar–Kullback–Leibler–Pinsker inequality and (44),

$$\|P_k \mu^{(N)} - f_{\delta_k}^{\otimes k}\|_{L^1(\mathbb{R}^k)} \leq \sqrt{2H(P_k \mu_{\delta_k}^{(N)} | f_{\delta_k}^{\otimes k})} < \frac{1}{k}$$

for all $N > N_k$. Therefore, for all $N > N_k$,

$$\left| \int_{\mathbb{R}^k} \phi dP_k \mu^{(N)} - \int_{\mathbb{R}^k} \phi f_{\delta_k}^{\otimes k} dv \right| < \frac{\|\phi\|_\infty}{k},$$

while trivial estimates show that

$$\lim_{\delta \rightarrow 0} \left| \int_{\mathbb{R}^k} \phi f_{\delta_k}^{\otimes k} dv - \int_{\mathbb{R}^k} \phi f^{\otimes k} dv \right| = 0.$$

Finally, the fact that (iii) holds follows easily from (43). □

7. Entropy production bounds. In this section, we prove Theorem 3. We first construct initial data f for the Boltzmann–Kac equation that has low entropy production. We then show that this implies that the fdv -chaotic family of initial data for the Kac master equation also has low entropy production.

Given two probability densities f and g on the line \mathbb{R} , define

$$(f \circ g)(v) = \frac{1}{2\pi} \int_{\mathbb{R}} \int_0^{2\pi} f(\cos(\theta)v - \sin(\theta)v_*) g(\sin(\theta)v + \cos(\theta)v_*) d\theta dv_* .$$

The density $f \circ g$ is called the *Wild convolution* of f and g , and using it we may write the Boltzmann–Kac equation in the compact form

$$\frac{\partial}{\partial t} f = f \circ f - f . \quad (45)$$

Any rescaling of $\gamma(v)$ is an equilibrium solution of this equation: For any $a > 0$, define

$$M_a = (2\pi a)^{-1/2} \exp(-v^2/2a) = a^{-1/2} \gamma(a^{-1/2}v) .$$

There are the so-called *Maxwellian densities*, and one easily sees that for all $a > 0$, $M_a \circ M_a = M_a$, so that these are stationary solutions of (45).

For any zero mean, unit variance solution f of (45), the relative entropy with respect to γ satisfies

$$-\frac{d}{dt} H(f|\gamma) = \int_{\mathbb{R}} (-\log f)[f \circ f - f] dv . \quad (46)$$

The analog of the Boltzmann H -Theorem for the Boltzmann–Kac equation asserts that this quantity is strictly positive unless f is one of the Maxwellians, in which case it is zero.

Our first goal in this section is to construct, for each $c > 0$, zero mean, unit variance initial data f for which

$$\int_{\mathbb{R}} (-\log f)[f \circ f - f] dv < c H(f|\gamma) .$$

There is a very natural construction that has been exploited by Bobylev and Cercignani [4] in the case of the actual Boltzmann equation: *Use a superposition of two very different Maxwellians.* This is natural since each M_a is an equilibrium solution. (In the case of the actual Boltzmann equation there is an even larger class of equilibrium densities to work with since momentum is also conserved. Here, only centered Maxwellians are equilibrium solutions.)

Pick a small positive number δ , and define

$$f = (1 - \delta)M_a + \delta M_b \quad (47)$$

where

$$b = 1/(2\delta) \quad \text{and} \quad a = 1/(2(1 - \delta)) . \quad (48)$$

Then since $\int_{\mathbb{R}} M_a dv = 1$ and $\int_{\mathbb{R}} v^2 M_a dv = a$,

$$\int_{\mathbb{R}} v^2 (1 - \delta) M_a dv = \int_{\mathbb{R}} v^2 \delta M_b dv = \frac{1}{2} ,$$

so that each Maxwellian component contributes half of the energy, though for small δ , most of the mass is contained in the M_a component.

Proposition 23. For any $c > 0$ there is a probability distribution f on \mathbb{R} such that $\int v^2 f(v) dv = 1$ and

$$\frac{D(f)}{H(f|\gamma)} \leq c,$$

where $D(f) = \int_{\mathbb{R}} (-\log f) [f \circ f - f] dv$ is the entropy production for the Boltzmann–Kac equation. Moreover, f can be chosen smooth with finite moments of all orders, in fact a linear combination of Gaussian functions.

Proof of Proposition 23. Define $f = (1 - \delta)M_a + \delta M_b$ as above. We first show that $H(f|\gamma)$ is bounded away from 0 uniformly for all δ sufficiently small. In fact,

$$\liminf_{\delta \rightarrow 0} H(f|\gamma) \geq \frac{1}{4}. \quad (49)$$

To prove (49), we use the definition of f and the monotonicity of the logarithm:

$$\begin{aligned} \int_{\mathbb{R}} f \log \left(\frac{f}{\gamma} \right) dv &= (1 - \delta) \int_{\mathbb{R}} M_a \log \left(\frac{f}{\gamma} \right) dv + \delta \int_{\mathbb{R}} M_b \log \left(\frac{f}{\gamma} \right) dv \\ &\geq (1 - \delta) \int_{\mathbb{R}} M_a \log \left(\frac{(1 - \delta)M_a}{\gamma} \right) dv + \delta \int_{\mathbb{R}} M_b \log \left(\frac{\delta M_b}{\gamma} \right) dv \\ &= (1 - \delta) \log(1 - \delta) + (1 - \delta)H(M_a|\gamma) + \delta \log \delta + \delta H(M_b|\gamma). \end{aligned}$$

Now use $H(M_c|\gamma) = \frac{1}{2}(c - 1) - \frac{1}{2} \log c$, and (49) at once.

It remains to estimate the entropy production associated with f . First,

$$f \circ f = (1 - \delta)^2 M_a + \delta^2 M_b + 2\delta(1 - \delta)M_a \circ M_b.$$

Therefore,

$$\begin{aligned} &\int_{\mathbb{R}} (-\log f) [f \circ f - f] dv \\ &= \int_{\mathbb{R}} (-\log f) [(\delta^2 - \delta)M_a + (\delta^2 - \delta)M_b + 2\delta(1 - \delta)M_a \circ M_b] dv \\ &= (\delta - \delta^2) \int_{\mathbb{R}} (-\log f) [2M_a \circ M_b - (M_a + M_b)] dv. \end{aligned}$$

Next, use the fact that for all $\delta < 1$, $(-\log f) \geq 0$. Hence, we can simplify the above to

$$\int_{\mathbb{R}} (-\log f) [f \circ f - f] dv \leq 2\delta \int_{\mathbb{R}} (-\log f) M_a \circ M_b dv.$$

By monotonicity of the logarithm, $\log f \geq \log(\delta M_b)$ so that

$$(-\log f) \leq -\frac{1}{2}(3 \log \delta - \log \pi) + \delta v^2.$$

Hence we have

$$\int_{\mathbb{R}} (-\log f) [f \circ f - f] dv \leq 2\delta \left(-\frac{1}{2}(3 \log \delta - \log \pi) \right) + 2\delta^2.$$

Evidently, the leading term is $-3\delta \log \delta$. So

$$\lim_{\delta \rightarrow 0} \int_{\mathbb{R}} (-\log f) [f \circ f - f] dv = 0. \quad (50)$$

The combination of (49) and (50) implies Proposition 23 at once. \square

Remark 24. The densities f constructed above do not, in their dependence on δ , have uniformly bounded fourth moments. In fact, the ratio of $\int v^4 f dv$ to $\int v^4 M_1 dv$ tends to infinity like $1/\delta$.

Now we shall deduce Theorem 3 from Proposition 23:

Proof of Theorem 3. Let $c > 0$, and let f be defined by Proposition 23. Let further

$$F^N = [f^{\otimes N}]_{S^{N-1}(\sqrt{N})} .$$

Theorem 9 guarantees that

$$\frac{H(F^N | \sigma^N)}{N} \longrightarrow H(f | \gamma) .$$

So it suffices to establish

$$\lim_{N \rightarrow \infty} \frac{1}{N} \langle \log(F^N), L_N(F^N) \rangle_{L^2(S^{N-1}(\sqrt{N}))} = 2 \int_{\mathbb{R}} \log f [f \circ f - f] dv . \quad (51)$$

To prove (51), we first write

$$\log F^N = -\log Z_N(f, \sqrt{N}) + \sum_{k=1}^N \log f(v_k)$$

and hence $\langle L_N(F^N), \log(F^N) \rangle_{L^2(S^{N-1}(\sqrt{N}))}$ is given by

$$\begin{aligned} & \int_{S^{N-1}(\sqrt{N})} \left[-\log(Z_N(f, \sqrt{N})) + \sum_{k=1}^N \log f(v_k) \right] N (QF^N - F^N) d\sigma^N \\ &= \int_{S^{N-1}(\sqrt{N})} \left[\sum_{k=1}^N \log f(v_k) \right] N (QF^N - F^N) d\sigma^N \\ &= \sum_{k=1}^N \sum_{i < j}^N \int_{S^{N-1}(\sqrt{N})} \log f(v_k) \frac{2}{N-1} \\ & \times \left(\frac{1}{2\pi} \int_0^{2\pi} (f^{\otimes N}(R_{i,j,\theta} V) - f^{\otimes N}(V)) d\theta \right) \frac{1}{Z_N(f, \sqrt{N})} d\sigma^N . \end{aligned} \quad (52)$$

By the invariance of σ^N under rotations, the integral over $S^{N-1}(\sqrt{N})$ vanished unless $k = i$ or $k = j$. By the permutation symmetry, we may set $k = 1$, and account for the sum over k by multiplying by N . There are $N-1$ pairs of which 1 is a member. We may set $i, j = 1, 2$, and account for the sum over pairs by multiplying by $N-1$. We are then left with

$$2N \int_{S^{N-1}(\sqrt{N})} \log f(v_1) \left(\frac{1}{2\pi} \int_0^{2\pi} (f(v'_1) f(v'_2) - f(v_1) f(v_2)) d\theta \right) \times \frac{1}{Z_N(f, \sqrt{N})} \prod_{j=3}^N f(v_j) d\sigma^N \quad (53)$$

where

$$v'_1 = \cos(\theta)v_1 + \sin(\theta)v_2 \quad \text{and} \quad v'_2 = -\sin(\theta)v_1 + \cos(\theta)v_2 .$$

Since f is a linear combination of Maxwellian densities, there is a constant C such that $|\log f(v)| \leq C(1 + v^2)$. Then again since f is a linear combination of Maxwellian densities,

$$\left| \log f(v) (f(v'_1)f(v'_2) - f(v_1)f(v_2)) \right|$$

is bounded, and in fact has Gaussian decay.

The proof will be completed by showing that

$$\lim_{N \rightarrow \infty} P_2 \left(\frac{1}{Z_N(f, \sqrt{N})} \prod_{j=3}^N f(v_j) d\sigma^N \right) = dv_1 dv_2 .$$

Since, with $s^2 = v_1^2 + v_2^2$,

$$P_2 \left(\frac{1}{Z_N(f, \sqrt{N})} \prod_{j=3}^N f(v_j) d\sigma^N \right) = \frac{Z_{N-2}(f, \sqrt{N-s^2})}{Z_N(f, \sqrt{N})} P_2(d\sigma_N) ,$$

it remains only to show that

$$\lim_{N \rightarrow \infty} \frac{Z_{N-2}(f, \sqrt{N-s^2})}{Z_N(f, \sqrt{N})} = 2\pi e^{-(v_1^2+v_2^2)/2} \quad (54)$$

since it is well known, and easily follows from (7), that

$$\lim_{N \rightarrow \infty} P_2(d\sigma^N) = \frac{1}{2\pi} e^{-(v_1^2+v_2^2)/2} dv_1 dv_2 .$$

However, (54) easily follows from Theorem 4 and Stirling's formula. \square

Remark 25. We have taken advantage of Maxwellian bounds on f to shorten the proof, but a similar result could be obtained in the same way for more general densities f by arguing as in the proof of Theorem 20. A more challenging problem would be to prove an analog of (51) for a more general class of chaotic data than conditioned tensor products.

Appendix A. An entropic Local Central Limit Theorem. The terminology ‘‘Local Central Limit Theorem’’ is used to designate a version of the Central Limit Theorem in which the conclusion is strengthened from weak convergence of the law to locally uniform pointwise convergence of the densities [14].

As recalled in the introduction, such a theorem can only hold if the common law of the independent random variables has a density f that satisfies certain regularity hypotheses — in particular, f is usually required to belong to L^p for some $p > 1$.

Of course the rate of pointwise convergence depends on the regularity of f . In this paper, we require precise, quantitative information on the rate, and the version of the Local Central Limit Theorem that we prove here provides this.

There is a remarkable feature that emerges: When an L^p bound is imposed on f , then the asymptotic rate of (pointwise) convergence of the densities $\sqrt{N} f^{*N}(\sqrt{N}x)$ to the Gaussian distribution can be estimated in terms of only the relative entropy $H(f|\gamma)$, even if the assumption $H(f|\gamma) < \infty$ alone is *not* sufficient to ensure the convergence! Of course, the L^p bound on f enters the estimates of convergence, *but only in determining how large N must be before the universal rate estimates governed by $H(f|\gamma)$ become valid.* For this reason, we refer to the result obtained here as an Entropic Local Central Limit Theorem.

In addition to the L^p bound on f , one requires a certain measure of localization of $f(v)v^2dv$, which is usually taken care of in the assumptions by assuming that

$$\int_{\mathbb{R}} f(v)v^{2+\epsilon}dv < \infty \quad (55)$$

for some $\epsilon > 0$.

So that we may include all finite energy initial data in our conclusions, we wish to avoid such an assumption. For this reason, we introduce the function

$$\chi(r) = \int_{|x| \geq \frac{1}{r}} |x|^2 f(x) dx .$$

This is a continuous function vanishing at $r = 0$, and the rate at which it vanishes as $r \rightarrow 0$ gives the sort of control that would be provided by (55). Indeed, under the assumption (55),

$$\chi(r) = \mathcal{O}(r^\epsilon) .$$

Before beginning the proof, we note that in this Appendix the state space is \mathbb{R}^k , where k is some positive integer. All the constants in our results may depend on k , but this dependence will not be explicitly recalled.

We start by some properties of the Fourier transform of probability densities. In the sequel, we use the following convention for the Fourier transform in \mathbb{R}^k :

$$\widehat{f}(\xi) = \int_{\mathbb{R}^k} e^{-2i\pi x\xi} f(x) dx .$$

Proposition 26. Let g be a probability density on \mathbb{R}^k , such that

$$\int_{\mathbb{R}^k} xg(x) dx = 0, \quad \int_{\mathbb{R}^k} (x \otimes x) g(x) dx = I_k, \quad \int_{\mathbb{R}^k} g \log g dx \leq H < +\infty,$$

where I_k is the $k \times k$ identity matrix. Let further χ be such that

$$\int_{|x| \geq \frac{1}{r}} |x|^2 g(x) dx \leq \chi(r),$$

where $\chi(r)$ goes to 0 as $r \rightarrow 0$. Then

(i) Given $\eta > 0$ there is $\alpha = \alpha(H, \eta) > 0$ such that

$$|\xi| \geq \eta \implies |\widehat{g}(\xi)| \leq 1 - \alpha.$$

(ii) There is a function $\varepsilon(\delta) = \varepsilon(H, \chi, \delta)$, going to 0 as $\delta \rightarrow 0$, such that

$$|\xi| \leq \delta \implies \left| \widehat{g}(\xi) - (1 - 2\pi^2|\xi|^2) \right| \leq \varepsilon(\delta)|\xi|^2.$$

(iii) There is $\alpha_0 = \alpha_0(H, \chi)$ such that

$$\forall \xi \in \mathbb{R}^k \quad |\widehat{g}(\xi)| \leq \max(1 - \pi^2|\xi|^2, 1 - \alpha_0).$$

Proof. First, it is clear that (iii) follows from (i) and (ii). For simplicity, we shall prove (i) and (ii) only in the case $k = 1$; the generalization to higher dimension does not bring in any major complication.

Let us prove (i). Let ξ be such that $|\xi| \geq \eta$, and let z be such that $\widehat{g}(\xi)e^{-2i\pi z\xi} = |\widehat{g}(\xi)|$. Then, with \Re standing for real part,

$$\begin{aligned} |\widehat{g}(\xi)| &= \Re[\widehat{g}(\xi)e^{-2i\pi z\xi}] \\ &= \Re\left(\int g(x)e^{-2i\pi(x+z)\xi} dx\right) \\ &= \int g(x) \cos[2\pi(x+z)\xi] dx. \end{aligned}$$

Let $R > 1$, to be chosen later. Write

$$\begin{aligned} |\widehat{g}(\xi)| &= \int g(x) dx - \int g(x)(1 - \cos[2\pi(x+z)\xi]) dx \\ &= 1 - \int g(x)(1 - \cos[2\pi(x+z)\xi]) dx \\ &\leq 1 - \int_{[-R,R]} g(x)(1 - \cos[2\pi(x+z)\xi]) dx. \end{aligned}$$

So it is sufficient to show that

$$\int_{[-R,R]} g(x)(1 - \cos[2\pi(x+z)\xi]) dx \geq \alpha > 0.$$

Let $\beta \in (0, 1/2)$, to be chosen later; define

$$B := \left\{ x \in [-R, R]; 1 - \cos[2\pi(x+z)\xi] \leq \beta \right\}.$$

The point is to show that $\int_B g$ is small when β is small too. For this we shall show that B has small Lebesgue measure and use the entropy bound on g .

If x lies in B , then $|x| \leq R$, and there exists $n \in \mathbb{Z}$ such that

$$\left| x - \frac{n}{\xi} \right| \leq \frac{\cos^{-1}(1 - \beta)}{2\pi|\xi|}.$$

So B consists of at most $2R|\xi| + 3$ intervals, with width $\cos^{-1}(1 - \beta)/(\pi|\xi|)$, which can be bounded by $\sqrt{2\beta}/(\pi|\xi|)$. So the Lebesgue measure $|B|$ of B can be estimated as follows:

$$\begin{aligned} |B| &\leq \left(\frac{2R|\xi| + 3}{\pi|\xi|} \right) \sqrt{2\beta} \leq \frac{2R}{\pi} \left(1 + \frac{1}{|\xi|} \right) \sqrt{2\beta} \\ &\leq \frac{2R}{\pi} \left(1 + \frac{1}{\eta} \right) \sqrt{2\beta}. \end{aligned} \tag{56}$$

Now define

$$\mu(dx) = \frac{g(x)1_{[-R,R]}(x) dx}{\int_{[-R,R]} g} \quad \nu(dx) = \frac{1_{[-R,R]}(x) dx}{2R}.$$

By direct computation,

$$\begin{aligned} H(\mu|\nu) &= \left(\int_{[-R,R]} g \right)^{-1} \int_{[-R,R]} g \log g - \log \left(\int_{[-R,R]} g \right) + \log(2R) \\ &\leq \int g |\log g| + \log(2R) - \log \left(1 - \frac{1}{R^2} \right), \end{aligned}$$

where we have used Chebyshev's inequality to get the bound on the last term: $\int_{|x|>R} g \leq (1/R^2) \int g x^2 dx \leq 1/R^2$. It is classical that $\int g |\log g|$ can be controlled by

$\int g \log g$ and $\int g x^2 dx = 1$: indeed, if γ stands for the standard gaussian distribution, then

$$\int_{g \leq 1} \left(g \log \frac{g}{\gamma} - g + \gamma \right) \geq 0,$$

so

$$\int_{g \leq 1} g \log g \geq \int_{g \leq 1} g \log \gamma + \int_{g \leq 1} g - \int_{g \leq 1} \gamma;$$

replacing $\log \gamma$ by its explicit expression, we obtain

$$\begin{aligned} \int_{g \leq 1} g \log g &\geq - \int_{g \leq 1} \frac{x^2}{2} g(x) dx + \left(1 - \frac{\log(2\pi)}{2} \right) \int_{g \leq 1} g - \int_{g \leq 1} \gamma \\ &\geq - \left(1 + \frac{\log(2\pi)}{2} \right). \end{aligned}$$

The desired bound follows, since

$$\int g |\log g| = \int g \log g - 2 \int_{g \leq 1} g \log g.$$

To summarize: there is an explicit bound

$$H(\mu|\nu) \leq h(H, R). \quad (57)$$

On the other hand, it follows from (56) that

$$\nu[B] \leq \frac{1}{\pi} \left(1 + \frac{1}{\eta} \right) \sqrt{2\beta}.$$

Combining this with (57) and a general entropy inequality, we find

$$\mu[B] \leq \frac{2H(\mu|\nu)}{\log \left(1 + \frac{H(\mu|\nu)}{\nu[B]} \right)} \leq \frac{2h}{\log \left(1 + \frac{\pi h}{\left(1 + \frac{1}{\eta} \right) \sqrt{2\beta}} \right)}.$$

So if H and η are given, there is a function $m(\beta)$, going to 0 as $\beta \rightarrow 0$ and depending only on H and η , such that $\mu[B] \leq m(\beta)$.

The desired conclusion follows easily:

$$\begin{aligned} \int_{[-R, R]} g(x) \left(1 - \cos[2\pi(x+z)\xi] \right) dx &\geq \beta \int_{[-R, R]} g(x) dx \\ &\geq \beta \left(\int_{[-R, R]} g \right) \mu[\mathbb{R} \setminus [-R, R]] \\ &= \beta \left(1 - \int_{|x| > R} g \right) (1 - \mu[[-R, R]]) \\ &\geq \beta \left(1 - \frac{1}{R^2} \right) \left(1 - \frac{2h}{\log \left(1 + \frac{\pi h}{\left(1 + \frac{1}{\eta} \right) \sqrt{2\beta}} \right)} \right). \end{aligned}$$

This establishes (i) with

$$\alpha := \sup_{\beta, R} \beta \left(1 - \frac{1}{R^2} \right) \left(1 - \frac{2h}{\log \left(1 + \frac{\pi h}{\left(1 + \frac{1}{\eta} \right) \sqrt{2\beta}} \right)} \right).$$

To get a lower bound on α , one may choose for instance

$$R = 2, \quad \beta = e^{-8h} \left(\frac{(\pi h)^2}{2(1 + \eta^{-1})^2} \right),$$

then one finds

$$\alpha \geq \frac{3}{8} e^{-8h} \left(\frac{(\pi h)^2}{2(1 + \eta^{-1})^2} \right).$$

Now let us prove (ii). Assume for instance that $\xi > 0$. By Taylor formula,

$$\begin{aligned} e^{-2i\pi x \xi} &= 1 - 2i\pi x \xi - 4\pi^2 x^2 \int_0^\xi (\xi - \zeta) e^{-2i\pi x \zeta} d\zeta \\ &= 1 - 2i\pi x \xi - 4\pi^2 x^2 \left(\frac{\xi^2}{2} \right) + 4\pi^2 x^2 \int_0^\xi (\xi - \zeta) (1 - e^{-2i\pi x \zeta}) d\zeta. \end{aligned}$$

So for $|\xi| \leq \eta$, one has

$$\left| \widehat{g}(\xi) - (1 - 2\pi^2 \xi^2) \right| \leq \varepsilon \xi^2,$$

with

$$\begin{aligned} \varepsilon &= \frac{4\pi^2}{\xi^2} \left| \int_0^\xi (\xi - \zeta) ([1 - e^{-2i\pi x \zeta}] x^2 g(x) dx) d\zeta \right| \\ &\leq \frac{4\pi^2}{\xi^2} \left(\int_0^\xi (\xi - \zeta) d\zeta \right) \left(\sup_{|\xi| \leq \eta} \int |1 - e^{-2i\pi x \zeta}| x^2 g(x) dx \right) \\ &= 4\pi^2 \sup_{|\zeta| \leq \eta} \int |\sin(\pi \zeta x)| x^2 g(x) dx. \end{aligned}$$

For $|x| \leq 1/r$ and $|\zeta| \leq \eta$, one has $|\sin(2\pi \zeta x)| \leq |2\pi \zeta x| \leq 2\pi \eta / r$; on the other hand, for $|x| \geq 1/r$, we can use the trivial bound $|\sin(\pi \zeta x)| \leq 1$. So

$$\int |\sin(\pi \zeta x)| x^2 g(x) dx \leq \frac{2\pi \eta}{r} + \int_{|x| \geq 1/r} x^2 g(x) dx \leq \frac{2\pi \eta}{r} + \chi(r).$$

In conclusion, $\varepsilon \leq \inf_{r > 0} \left[\frac{2\pi \eta}{r} + \chi(r) \right]$, and the right-hand side goes to 0 as $\eta \rightarrow 0$.

This proves (ii). \square

Now we can proceed with the main result of this Appendix.

Theorem 27 (Local Central Limit Theorem). Let g be a probability distribution on \mathbb{R}^k , satisfying

$$\int_{\mathbb{R}^k} g(x) x dx = 0, \quad \int_{\mathbb{R}^k} g(x) (x \otimes x) dx = I_k, \quad \int g \log g dx \leq H,$$

where I_k is the $k \times k$ identity matrix. Let χ be such that $\chi(r) \rightarrow 0$ as $r \rightarrow 0$, and $\int_{|x| \geq \frac{1}{r}} g(x) dx \leq \chi(r)$. Let further $g_N(x) = \sqrt{N}^{-k} g^{*N}(\sqrt{N}x)$, for any positive integer N . Then:

(i) If $g \in L^p(\mathbb{R}^k)$, $1 < p < \infty$, then g_N is continuous for $N \geq p' = p/(p-1)$; and for any $\delta > 0$ there is $\alpha = \alpha(\chi, H, \delta) > 0$ and $\bar{\varepsilon} = \bar{\varepsilon}(\chi, H, \delta) > 0$ such that, given k ,

χ and H , $\bar{\varepsilon}(\delta) \xrightarrow{\delta \rightarrow 0} 0$ and, for $N \geq p'$,

$$\sup_{x \in \mathbb{R}^k} |g_N(x) - \gamma(x)| \leq \sqrt{N}(1 - \alpha)^{N-p'} \|g\|_{L^p}^{p'} + k \frac{e^{-2\pi^2 N \delta^2}}{\sqrt{N} \delta} + \bar{\varepsilon}(\delta),$$

where γ stands for the standard Gaussian distribution. In particular, $\sup |g_N - \gamma|$ goes to 0 as $N \rightarrow \infty$, and there is an upper bound on the rate of convergence which only depends on $k, p, \|g\|_{L^p}, \chi$ and H .

(ii) Given χ and H there is a function $\lambda(N)$, going to 0 as $N \rightarrow \infty$, such that if g lies in L^p for some $p \in (1, \infty)$, then there is $N_0 = N_0(\chi, H, p, \|g\|_{L^p})$ with

$$N \geq N_0 \implies \sup_{x \in \mathbb{R}^k} |g_N - \gamma| \leq \lambda(N).$$

Proof of Theorem 27. First, it follows from Young's convolution inequality that $g^{*(N-1)}$ lies in $L^{p'}(\mathbb{R}^k)$; then its convolution product with g is continuous.

By properties of the Fourier transform,

$$\widehat{g}_N(\xi) = \widehat{g}\left(\frac{\xi}{\sqrt{N}}\right)^N.$$

Without loss of generality, assume $p \leq 2$; then $\widehat{g}(\cdot/\sqrt{N})$ lies in $L^{p'} \cap L^\infty$ by the Hausdorff-Young inequality, so \widehat{g}_N lies in L^1 for $N \geq p'$. Then the Fourier inversion formula applies:

$$g_N(x) = \int_{\mathbb{R}^k} \widehat{g}_N(\xi) e^{2i\pi x \cdot \xi} d\xi.$$

In particular, for any $x \in \mathbb{R}$,

$$|g_N(x) - \gamma(x)| = \left| \int_{\mathbb{R}^k} (\widehat{g}_N(\xi) - \widehat{\gamma}(\xi)) e^{2i\pi x \cdot \xi} d\xi \right| \leq \int |\widehat{g}_N - \widehat{\gamma}|.$$

We separate between low and high frequencies, according to some threshold $\delta\sqrt{N}$, to choose later:

$$\sup_{x \in \mathbb{R}^k} |g_N(x) - \gamma(x)| \leq \int_{|\xi| > \delta\sqrt{N}} |\widehat{g}_N| + \int_{|\xi| > \delta\sqrt{N}} |\widehat{\gamma}| + \int_{|\xi| \leq \delta\sqrt{N}} |\widehat{g}_N - \widehat{\gamma}|. \quad (58)$$

To estimate the first term in the right-hand side of (58), we use Proposition 26 (i): there is $\alpha = \alpha(\delta, \chi, H) > 0$ such that $|\xi| \geq \eta \implies |\widehat{g}(\xi)| \leq 1 - \alpha$; so

$$\begin{aligned} \int_{|\xi| > \delta\sqrt{N}} |\widehat{g}_N| &= \int_{|\xi| > \delta\sqrt{N}} \left| \widehat{g}\left(\frac{\xi}{\sqrt{N}}\right) \right|^N d\xi = \sqrt{N}^k \int_{|\xi| > \delta} |\widehat{g}|^N d\xi \\ &\leq \sqrt{N}^k (1 - \alpha)^{N-p'} \int |\widehat{g}|^{p'}. \end{aligned}$$

Combining this with the Hausdorff-Young inequality, we find

$$\int_{|\xi| > \delta\sqrt{N}} |\widehat{g}_N| \leq \sqrt{N}^k (1 - \alpha)^{N-p'} \|g\|_{L^p}^{p'}. \quad (59)$$

The second term in the right-hand side of (58) can be bounded by an explicit estimate:

$$\int_{|\xi| > \delta\sqrt{N}} |\widehat{\gamma}| \leq k \frac{e^{-2\pi^2 N \delta^2}}{\delta\sqrt{N}}.$$

The third term in the right-hand side of (58) is a bit more subtle. On one hand, by a telescopic sum argument, since $|\widehat{g}| \leq 1$ and $|\widehat{\gamma}| \leq 1$, we have

$$|\widehat{g}_N(\xi) - \widehat{\gamma}(\xi)| = \left| \widehat{g} \left(\frac{\xi}{\sqrt{N}} \right)^N - \widehat{\gamma} \left(\frac{\xi}{\sqrt{N}} \right)^N \right| \leq N \left| \widehat{g} \left(\frac{\xi}{\sqrt{N}} \right) - \widehat{\gamma} \left(\frac{\xi}{\sqrt{N}} \right) \right|.$$

Now we can apply Proposition 26 (ii), with ξ replaced by ξ/\sqrt{N} and $\varepsilon = \max(\varepsilon_g, \varepsilon_\gamma)$ where ε_g is the function appearing in the Proposition:

$$\begin{aligned} N \left| \widehat{g} \left(\frac{\xi}{\sqrt{N}} \right) - \widehat{\gamma} \left(\frac{\xi}{\sqrt{N}} \right) \right| &\leq N \left| \widehat{g} \left(\frac{\xi}{\sqrt{N}} \right) - \left(1 - \frac{2\pi^2|\xi|^2}{N} \right) \right| + \\ &N \left| \widehat{\gamma} \left(\frac{\xi}{\sqrt{N}} \right) - \left(1 - \frac{2\pi^2|\xi|^2}{N} \right) \right| \leq 2N\varepsilon(\delta) \left(\frac{\xi}{\sqrt{N}} \right)^2 = 2\varepsilon(\delta) \xi^2. \end{aligned} \quad (60)$$

Here ε is a function depending only on H and χ .

On the other hand, by Proposition 26 (iii),

$$|\widehat{g}_N(\xi) - \widehat{\gamma}(\xi)| \leq \left| \widehat{g} \left(\frac{\xi}{\sqrt{N}} \right) \right|^N + \widehat{\gamma}(\xi) \leq \max \left(1 - \frac{\pi^2|\xi|^2}{N}, 1 - \alpha_0 \right)^N + \widehat{\gamma}(\xi).$$

Thanks to the inequality $(1 - u/N)^N \leq e^{-u}$, we conclude that

$$\begin{aligned} |\widehat{g}_N(\xi) - \widehat{\gamma}(\xi)| &\leq \max \left(e^{-\pi^2|\xi|^2}, (1 - \alpha_0)^N \right) + \widehat{\gamma}(\xi) \\ &\leq 2 \max \left(e^{-\pi^2|\xi|^2}, (1 - \alpha_0)^N \right). \end{aligned} \quad (61)$$

By taking the geometric mean of (60) and (61), we obtain

$$|\widehat{g}(\xi) - \widehat{\gamma}(\xi)| \leq 2\sqrt{\varepsilon} |\xi| \max \left(e^{-\frac{\pi^2|\xi|^2}{2}}, (1 - \alpha_0)^{\frac{N}{2}} \right).$$

Then

$$\begin{aligned} \int_{|\xi| \leq \delta\sqrt{N}} |\widehat{g}(\xi) - \widehat{\gamma}(\xi)| &\leq \sqrt{2\varepsilon} \left(\int_{|\xi| \leq \delta\sqrt{N}} |\xi| e^{-\frac{\pi^2|\xi|^2}{2}} d\xi + (1 - \alpha_0)^{\frac{N}{2}} \int_{|\xi| \leq \delta\sqrt{N}} |\xi| d\xi \right) \\ &\leq \sqrt{2\varepsilon} \left(\int |\xi| e^{-\frac{\pi^2|\xi|^2}{2}} d\xi + (1 - \alpha_0)^{\frac{N}{2}} |S^{k-1}| \frac{(\delta\sqrt{N})^{k+1}}{k+1} \right) \\ &\leq \sqrt{2\varepsilon} \left(\int |\xi| e^{-\frac{\pi^2|\xi|^2}{2}} d\xi + C(k, \alpha_0) \right) =: \overline{\varepsilon}(\delta, \chi, H) \end{aligned}$$

(where $C(k, \alpha_0)$ is a constant which does not depend on N). This concludes the proof of (i).

To prove (ii), we let p vary with N in such a way that p' remains of the order of N ; for instance, $p' = N/2$ (for N large enough). Then, as N goes to infinity,

$$\log \|g\|_{L^p}^{p'} = \frac{1}{p-1} \log \int g^p \xrightarrow{p \rightarrow 1} \int g \log g.$$

So, when N is large enough, the first term in the right-hand side of (i) can be bounded by $(1 - \alpha)^{N/2} e^H$, which does not depend on the L^p norm of g . (But “large enough” here may depend on this norm!!) \square

Acknowledgments. E.C., M.C. and M.L. thank the ENS Lyon for hospitality on a visit during which this work was begun and substantially developed. E.C., M.L. and C.V. thank C.M.A.F. and the University of Lisbon for hospitality on a visit during which additional work on this paper was undertaken. We thank the referees for a careful reading of the paper, and helpful suggestions.

REFERENCES

- [1] BARDOS, C., GOLSE, F., AND LEVERMORE, D. Fluid dynamical limits of kinetic equations, I : Formal derivation. *J. Statist. Phys.* **63** (1991), 323–344.
- [2] BARTHE, F., CORDERO–ERAUSQUIN, D., AND MAUREY, B. Entropy of spherical marginals and related inequalities, *J. Math. Pures Appl.* **86** (2006), 89–99.
- [3] BEN AROUS, G., AND ZEITOUNI, O. Increasing propagation of chaos for mean field models. *Ann. Inst. H. Poincaré Probab. Statist.* **35**, 1 (1999), 85–102.
- [4] BOBYLEV, A. V., AND CERCIGNANI, C. On the rate of entropy production for the Boltzmann equation. *J. Statist. Phys.* **94**, 3-4 (1999), 603–618.
- [5] CAPUTO, P. On the spectral gap of the Kac walk and other binary collision processes *ALEA Lat. Am. J. Probab. Math. Stat.* **4**, (2008), 205–222.
- [6] CARLEMAN, T. Sur la solution de l'équation intégrodifférentielle de Boltzmann. *Acta. Math.* **60**, 3 (1933), 91–146.
- [7] CARLEN, E. A. The rate of local equilibration in kinetic theory. *Contemporary Mathematics*, **437** (2007), 71–88.
- [8] CARLEN, E. A., CARVALHO, M. C. AND LOSS, M. Many–body aspects of approach to equilibrium. In *Journées Équations aux dérivées partielles*, N. D. et. al., Ed. Univ. of Nantes, Nantes, 2000, pp. 165–185.
- [9] CARLEN, E. A., CARVALHO, M. AND LOSS, M. Determination of the spectral gap for Kac's master equation and related stochastic evolution. *Acta Math.* **191** pp 1–54, 2003
- [10] CARLEN, E. A., LIEB, E. AND LOSS, M. A sharp analog of Young's inequality on S^N and related entropy inequalities. *Jour. Geom. Analysis* **14**, 3 (2004), 487–520.
- [11] CARLEN, E. A., GERNIMO, J. AND LOSS, M. Determination of the spectral gap in the Kac model for physical momentum and energy conserving collisions *To appear in SIAM Jour. on Math. Analysis*, 2008
- [12] CARLEN, E. A., AND LU, X. Fast and slow convergence to equilibrium Maxwellian molecules via Wild sums. *Jour. Stat. Phys.*, **112** no. 1/2, (2003) pp. 59–134
- [13] DONSKER, M. D., AND VARADHAN, S. R. S. Asymptotic evaluation of certain Markov processes for large time, I, *Comm. Pure Appl. Math.*, **28** (1975) pp. 1–47
- [14] FELLER, W. *An introduction to probability theory and its applications. Vol. II.* Second edition. John Wiley & Sons Inc., New York, 1971.
- [15] FOATA, D., AND FUCHS, A. *Calcul des probabilités.* Dunod, Paris, 1986.
- [16] GOLSE, F., AND SAINT-RAYMOND, L. The Navier-Stokes limit of the Boltzmann equation for bounded collision kernels. *Invent. Math.* **155**, 1 (2004), 81–161.
- [17] JANVRESSE, E. Spectral gap for Kac's model of Boltzmann equation. *Ann. Probab.* **29**, 1 (2001), 288–304.
- [18] KAC, M. Foundations of kinetic theory. Proc. 3rd Berkeley Symp. Math. Stat. Prob., J. Neyman, ed. Univ. of California, vol. 3, 171–197, 1956.
- [19] KOSYGINA, E. The behavior of the specific entropy in the hydrodynamic scaling limit. *Ann. Probab.* **29**, 3 (2001), 1086–1110.
- [20] LOTT, J., AND VILLANI, C. Ricci curvature for metric-measure spaces via optimal transport. *Ann. of Math.* **169**, 3 (2009), 903–991.
- [21] MAJOR, P. Series of problems in probability theory. The central limit theorem and Fourier analysis, I. Available online at www.renyi.hu/major/probability/central.html.
- [22] MASLEN, D. The eigenvalues of Kac's master equation. *Math. Zeit.* **243** (2003), 291–331.
- [23] MEHLER, F. G. Ueber die Entwicklung einer Function von beliebig vielen Variablen nach Laplaschen Functionen höherer Ordnung. *Crelle's Journal* **66** (1866), 161–176.
- [24] SZNITMAN, A.-S. Topics in propagation of chaos. In *École d'Été de Probabilités de Saint-Flour XIX—1989*. Springer, Berlin, 1991, pp. 165–251.
- [25] VILLANI, C. A review of mathematical topics in collisional kinetic theory. In *Handbook of mathematical fluid dynamics, Vol. I.* North-Holland, Amsterdam, 2002, pp. 71–305.

- [26] VILLANI, C. Cercignani's conjecture is sometimes true and always almost true. *Comm. Math. Phys.* 234, 3 (2003), 455–490.

Received xxxx 20xx; revised xxxx 20xx.

E-mail address: carlen@math.rutgers.edu

E-mail address: mcarvalh@cii.fc.ul.pt

E-mail address: leroux@hil.t.u-tokyo.ac.jp

E-mail address: loss@math.gatech.edu

E-mail address: cvillani@umpa.ens-lyon.fr