Commun. Math. Phys. 363, 847–875 (2018)
Digital Object Identifier (DOI) https://doi.org/10.1007/s00220-018-3263-0

Communications in
Mathematical Physics

CrossMark

Entropy Decay for the Kac Evolution

Federico Bonetto[1], Alissa Geisinger[2], Michael Loss[1], Tobias Ried[3]

[1] School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332, USA.
E-mail: bonetto@math.gatech.edu; loss@math.gatech.edu
[2] Universität Tübingen, Fachbereich Mathematik, Auf der Morgenstelle 10, 72076 Tübingen, Germany.
E-mail: alissa.geisinger@uni-tuebingen.de
[3] Institute for Analysis, Karlsruhe Institute of Technology (KIT), Englerstraße 2, 76131 Karlsruhe, Germany.
E-mail: tobias.ried@kit.edu

Received: 31 July 2017 / Accepted: 5 August 2018
Published online: 21 September 2018 –

Abstract: We consider solutions to the Kac master equation for initial conditions where N particles are in a thermal equilibrium and $M \leq N$ particles are out of equilibrium. We show that such solutions have exponential decay in entropy relative to the thermal state. More precisely, the decay is exponential in time with an explicit rate that is essentially independent on the particle number. This is in marked contrast to previous results which show that the entropy production for arbitrary initial conditions is inversely proportional to the particle number. The proof relies on Nelson's hypercontractive estimate and the geometric form of the Brascamp–Lieb inequalities due to Franck Barthe. Similar results hold for the Kac–Boltzmann equation with uniform scattering cross sections.

1. Introduction

Among the models describing a gas of interacting particles, the Kac master equation [21], due to its simplicity, occupies a special place. It is useful in illuminating various issues in kinetic theory, e.g., providing a reasonably satisfactory derivation of the spatially homogeneous Boltzmann equation and giving a mathematical framework for investigating the approach to equilibrium. These issues were, in fact, the motivation for Kac's original work. Although it does not have a foundation in Hamiltonian mechanics, the Kac master equation is based on simple probabilistic principles and yields a linear evolution equation for the velocity distribution for N particles undergoing collisions. It is in this context that Kac invented the notion of propagation of chaos and used this notion to derive the spatially homogeneous, non-linear Kac–Boltzmann equation. Kac also suggested various avenues to investigate the long time behaviour of the evolution and its approach to equilibrium as the number of particles, N, becomes large. He emphasized that this could be done in a quantitative way if one could show, e.g., that the gap of the generator is bounded below uniformly in N. This, known as Kac's conjecture, was proved by Élise Janvresse in [20] and, as a further sign of the simplicity of the model, the gap was computed explicitly in [9,10], see also [24]. One of the problems in using the gap is that the approach to equilibrium is measured in terms of an L^2 distance. While this does seem to be a natural way to look at this problem, the size of the L^2 norm of approximately independent probability distributions increases exponentially with the size of the system. Thus, the half life of the L^2 norm is of order N.

A natural measure is, of course, given by the entropy, which is extensive, i.e, proportional to N. There has not been much success in proving exponential decay of the entropy with good rates. In [29] Cédric Villani showed that the entropy decays exponentially, albeit with a rate that is bounded below by a quantity that is inversely proportional to N. This estimate was complemented by Amit Einav [14], who gave an example of a state that has entropy production essentially of order $1/N$. He chose a state in which most of the energy is concentrated in a few particles while most of the others have very little energy. One might surmise, based on physical intuition, that this state is physically very improbable and still has low entropy production because most of the particles are in some sort of equilibrium. This intuition can be made rigorous, see [14], although by a quite difficult computation. One should add that low entropy production does not preclude exponential decay in entropy, i.e., large entropy production for the initial state might not be necessary for an exponential decay rate for the entropy.

A breakthrough was achieved by Mischler and Mouhot in [25,26]. They undertook a general investigation of the Kac program for gases of hard spheres and true Maxwellian molecules in three dimensions. Among the results of Mischler and Mouhot is a proof that these systems relax towards equilibrium in relative entropy as well as in Wasserstein distance with a rate that is independent of the particle number. As expected, they achieve this not for any initial condition, but rather for a natural class of chaotic states. The rate of relaxation is, however, polynomial in time.

To summarize, there is so far no mathematical evidence that the entropy in the Kac model in general decays exponentially with a rate that is independent of N and physical intuition suggests that for highly "improbable" states, such as the one used by Einav, this cannot be expected. One can restrict the class of initial conditions by considering chaotic states as done by Mischler and Mouhot, which shifts the problem of finding suitable initial conditions for proving exponential decay to the level of the non-linear Boltzmann equation.

In this paper we take a different approach, one which is based on the idea of coupling a system of particles to a reservoir. The simplest such model is a finite system interacting with an *infinite* reservoir which is modeled by a thermostat. Recall from [7] the master equation of M particles with velocities $\mathbf{v} = (v_1, v_2, \ldots, v_M)$ interacting with a thermostat at temperature $1/\beta$,

$$\frac{\partial f}{\partial t} = \mathcal{L}_T f, \;\; f(\mathbf{v}, 0) = f_0(\mathbf{v}). \tag{1}$$

The operator $\mathcal{L}_T$ is given by

$$\mathcal{L}_T f = \mu \sum_{j=1}^{M} (B_j - I) f,$$

where

$$B_j[f](\mathbf{v}) := \int_{\mathbb{R}} \mathrm{d}w \int_{-\pi}^{\pi} \rho(\theta)\, \mathrm{d}\theta \sqrt{\frac{\beta}{2\pi}}\, e^{-\beta w_j^*(\theta)^2/2} f(\mathbf{v}_j(\theta, w)),$$

$$\mathbf{v}_j(\theta, w) = (v_1, \ldots, v_j \cos(\theta) + w \sin(\theta), \ldots, v_M) \text{ and}$$
$$w_j^*(\theta) = -v_j \sin(\theta) + w \cos(\theta),$$

and ρ is a probability distribution on $[-\pi, \pi]$. Thus, $B_j[f](\mathbf{v})$ describes the effect of a collision between particle j in the system and a particle in the thermostat. One way of thinking about this process is that the thermostat contains infinitely many particle each of which has a probability $\sqrt{\frac{\beta}{2\pi}}\, e^{-\beta v^2/2}$ of having velocity v. When a collision takes place a particle is randomly selected from the thermostat and discarded after the collision with particle j from the system. This ensures that the thermostat stays in equilibrium. The interaction times with the thermostat are given by a Poisson process whose intensity μ is chosen so that the average time between two successive interactions of a given particle with the thermostat is independent of the number of particles in the system. For the case where $\rho(\theta) = (2\pi)^{-1}$, the entropy decays exponentially fast. In fact, abbreviating $\sqrt{\beta/(2\pi)}e^{-\beta/2\mathbf{v}^2} = \Gamma_\beta(\mathbf{v})$, we know from [7], that

$$S(f(\cdot, t)) := \int_{\mathbb{R}^M} f(\mathbf{v}, t) \log\left(\frac{f(\mathbf{v}, t)}{\Gamma_\beta(\mathbf{v})}\right) d\mathbf{v} \le e^{-\mu t/2} S(f_0).$$

Thus, one might guess that if a “small" *system* of M particles out of equilibrium interacts with a *reservoir*, that is a large but *finite* system of $N \ge M$ particles in thermal equilibrium, then the entropy decays exponentially fast in time. This intuition is also supported by the results in [6]. There it was shown that if the thermostat is replaced by a large but finite reservoir *initially* in thermal equilibrium, this evolution is close to the evolution given by the thermostat. This result holds in various norms and, in particular, it is *uniform* in time. We would like to emphasize that the reservoir, because it is finite, will not stay in thermal equilibrium as time progresses, nevertheless it will not veer far from it.

The precise description of a system of M particles interacting with a reservoir of N particles is the following. We consider probability distributions $F : \mathbb{R}^{M+N} \to \mathbb{R}_+$ and write $F(\mathbf{v}, \mathbf{w})$ where $\mathbf{v} = (v_1, \ldots, v_M)$ describes the particles in the small system, whereas $\mathbf{w} = (w_{M+1}, \ldots, w_{N+M})$ describes the particles in the large system. The Kac master equation is given by

$$\frac{\partial F}{\partial t} = \mathcal{L}F, \;\; F(\mathbf{v}, \mathbf{w}, 0) = F_0(\mathbf{v}, \mathbf{w}) = f_0(\mathbf{v})e^{-\pi|\mathbf{w}|^2}, \tag{2}$$

where

$$\begin{aligned}\mathcal{L} =& \frac{\lambda_S}{M-1} \sum_{1\le i<j\le M} \left(R_{ij} - I\right) + \frac{\lambda_R}{N-1} \sum_{M<i<j\le N+M} \left(R_{ij} - I\right) \\ &+ \frac{\mu}{N} \sum_{i=1}^{M} \sum_{j=M+1}^{M+N} \left(R_{ij} - I\right),\end{aligned} \tag{3}$$

with

$$(R_{ij}F)(\mathbf{v}, \mathbf{w}) = \int_{-\pi}^{\pi} \rho(\theta)\, \mathrm{d}\theta\; F(r_{ij}(\theta)^{-1}(\mathbf{v}, \mathbf{w})),$$

and r_{ij} is given as follows. For $1 \le i < j \le M$ we have

$$r_{ij}(\theta)^{-1}(\mathbf{v}, \mathbf{w}) = (v_1, \ldots, v_i \cos\theta - v_j \sin\theta, \ldots, v_i \sin\theta + v_j \cos\theta, \ldots, v_M, \mathbf{w}), \tag{4}$$

while for $M + 1 \le i < j \le N$ we have

$$r_{ij}(\theta)^{-1}(\mathbf{v}, \mathbf{w}) = (\mathbf{v}, w_1, \ldots, w_i \cos\theta - w_j \sin\theta, \ldots, w_i \sin\theta + w_j \cos\theta, \ldots, w_N) \,. \tag{5}$$

Finally for $1 \le i \le M < j \le N$ we have

$$\begin{aligned} r_{ij}(\theta)^{-1}(\mathbf{v}, \mathbf{w}) = (&v_1, \ldots, v_i \cos\theta - w_j \sin\theta, \ldots, v_M, \\ &w_1, \ldots, v_i \sin\theta + w_j \cos\theta, \ldots, w_N) \,. \end{aligned} \tag{6}$$

We assume that the probability measure ρ is smooth and satisfies

$$\int_{-\pi}^{\pi} \rho(\theta)\, \mathrm{d}\theta \; \sin\theta \cos\theta = 0. \tag{7}$$

In particular, we do *not* require $\mathcal{L}$ to be self-adjoint on $L^2(\mathbb{R}^{N+M})$, a condition known as *microscopic reversibility*. The initial state of the reservoir is assumed to be a thermal equilibrium state and we have chosen units in which the inverse temperature $\beta = 2\pi$. Note that λ_S is the rate at which one particle from the system S will scatter with any other particle in the system S itself and similarly for λ_R. Likewise, μ is the rate at which a single particle of the system S will scatter with any particle in the reservoir R. This is due to the factor $1/N$ in front of the last sum in (3). Note that the rate at which a particular particle from the reservoir R will scatter with any particle in the system S is instead given by $\mu M/N$. Hence, when N is large compared to M this process is suppressed and one expects that the reservoir does not move far from its equilibrium. Indeed, it is shown in [6] that the solution of the master equation (3) stays close to the solution of a thermostated system in the Gabetta-Toscani-Wennberg metric,

$$d_{GTW}(F, G) := \sup_{k \neq 0} \frac{|\widehat{F}(k) - \widehat{G}(k)|}{|k|^2},$$

where $\widehat{F}$ denotes the Fourier transform of F, see [16]. More precisely, with the initial conditions (1) and (2), it was shown that

$$d_{GTW}(f(\mathbf{v}, t)e^{-\pi|\mathbf{w}|^2}, F(\mathbf{v}, \mathbf{w}, t)) \le C(f_0)\frac{M}{N},$$

where $C(f_0)$ is a constant that depends on the initial condition but is of order one. The distance varies inversely as N, the size of the reservoir and, moreover, this estimate holds *uniformly* in time. For a detailed description of the results we refer the reader to [6]. From this result and the fact that the entropy of the system interacting with a thermostat decays exponentially in time, one might surmise that the entropy of the system interacting with a finite reservoir also decays exponentially fast in time. In fact we shall show this to be true if we consider the entropy relative to the thermal state.

2. Results

For the solution of the master equation (2) we write

$$F(\mathbf{v}, \mathbf{w}, t) = (e^{\mathcal{L}t} F_0)(\mathbf{v}, \mathbf{w}). \tag{8}$$

This evolution preserves the energy and hence it is customary to consider it on $L^1(\mathbb{S}^{N+M-1}(\sqrt{N+M}))$ with the normalized surface measure. Likewise, under mild assumption on ρ it is easy to see that the evolution is ergodic on $L^1(\mathbb{S}^{N+M-1}(\sqrt{N+M}))$ in the sense that $e^{\mathcal{L}t} F_0 \to 1$ as $t \to \infty$ and 1 is the only normalized equilibrium state. So far our description matches the standard Kac master equation. For our purpose it will be convenient to describe the evolution as an evolution in $L^1(\mathbb{R}^{M+N})$ with Lebesgue measure. This space is fibered into spheres such that the integral of the solution over each of them is preserved, in fact $e^{\mathcal{L}t} F_0$ converges to the spherical average of F_0 taken over spheres in $\mathbb{R}^{M+N}$. As mentioned before, we choose the initial condition

$$F_0(\mathbf{v}, \mathbf{w}) = f_0(\mathbf{v}) e^{-\pi|\mathbf{w}|^2}. \tag{9}$$

Moreover, since we are only interested in the evolution of the system of M particles, we integrate over the velocities of the particles in the reservoir, i.e., we consider

$$f(\mathbf{v}, t) := \int_{\mathbb{R}^N} \left[e^{\mathcal{L}t} F_0 \right] (\mathbf{v}, \mathbf{w}) \, \mathrm{d}\mathbf{w} \tag{10}$$

and we call

$$S(f(\cdot, t)) := \int_{\mathbb{R}^M} f(\mathbf{v}, t) \log \left(\frac{f(\mathbf{v}, t)}{e^{-\pi|\mathbf{v}|^2}} \right) \mathrm{d}\mathbf{v},$$

the entropy of f *relative* to the thermal state $e^{-\pi|\mathbf{v}|^2}$. Our main result is the following theorem.

Theorem 1. *Let $N \geq M$ and let ρ be a probability distribution with an absolutely convergent Fourier series such that* (7) *holds. The entropy of f relative of to the thermal state $e^{-\pi|\mathbf{v}|^2}$ then satisfies*

$$S(f(\cdot, t)) \leq \left[\frac{M}{N+M} + \frac{N}{N+M} e^{-t\mu_\rho (N+M)/N} \right] S(f_0),$$

where

$$\mu_\rho = \mu \int_{-\pi}^{\pi} \rho(\theta) \, \mathrm{d}\theta \, \sin^2(\theta),$$

and f_0 is as introduced in (9).

Remark 1. 1. It should not come as a surprise that this entropy does not tend to zero as $t \to \infty$ since the entropy is measured relative to the thermal state which is a Gaussian and which is in general *not the equilibrium state*. The real question should be formulated in terms of the entropy relative to the equilibrium state which, as mentioned before is the spherical average of the initial condition. While it is easy to see that the entropy with respect to the equilibrium state tends to zero as $t \to \infty$ we do not know how to quantify this fact. Our method uses inequalities that are saturated by Gaussians which is the chief reason for obtaining such a clean result for thermal states. Equivalently, instead of considering the entropy with respect to the equilibrium state

on $L^1(\mathbb{R}^{M+N})$, one may treat the Kac master equation on $L^1(\mathbb{S}^{N+M-1}(\sqrt{N+M}))$ with an initial condition F_0 that is invariant under all rotations that fix the first M coordinates. The problem is to find the rate at which the entropy of the evolved state *relative* to the equilibrium state 1 decays to zero. As we said before, we do not know how to compute this rate.
2. The decay rate is universal in the sense that it only depends on μ and the distribution ρ. The intra-particle interactions in the system and in the reservoir do not seem to matter.
3. The statement of the theorem becomes particularly simple as $N \to \infty$. This corresponds to the thermostat problem treated in [7] with the exact same decay rate. It is known that for the thermostat the decay rate is optimal, see [28], and hence the decay rate here is optimal as well.
4. Although we assume that ρ is smooth, our result also holds for the case where ρ is a finite sum of Dirac measures. In particular Theorem 1 also holds if ρ is a delta measure that has its mass at the angles $\theta = \pm\pi/2$, that is, our result does not depend on ergodicity of the evolution.

As a consequence of (2) in Remark 1, one obtains a result for the standard Kac model. Recall that the generator of the standard Kac model is given by[1]

$$\mathcal{L}_{\text{cl}} = \frac{2}{N+M-1} \sum_{1\le i<j\le N+M} (R_{ij} - I).$$

We may somewhat arbitrarily split the variables into two groups $(v_1, \dots, v_M)$ and $(w_{M+1}, \dots, w_{M+N})$. Splitting the generator accordingly,

$$\mathcal{L}_{\text{cl}} = \frac{2}{N+M-1} \sum_{1\le i<j\le M} (R_{ij} - I) + \frac{2}{N+M-1} \sum_{M+1\le i<j\le N+M} (R_{ij} - I)$$
$$+ \frac{2}{N+M-1} \sum_{i=1}^{M} \sum_{j=M+1}^{N+M} (R_{ij} - I),$$

we see that the standard Kac model can be cast in the from (3) by setting

$$\lambda_S = \frac{2(M-1)}{N+M-1}\ , \ \lambda_R = \frac{2(N-1)}{N+M-1} \text{ and } \mu = \frac{2N}{N+M-1}.$$

Hence, we obtain the following Corollary.

Corollary 1. *Let $N \ge M$ and consider the time evolution defined by $\mathcal{L}_{\text{cl}}$ with initial condition* (9). *Assume that the function f_0 in the initial condition has finite entropy. The entropy of the function*

$$f(\mathbf{v}, t) := \int_{\mathbb{R}^N} \left[e^{\mathcal{L}_{\text{cl}} t} F_0 \right] (\mathbf{v}, \mathbf{w})\, \mathrm{d}\mathbf{w}$$

relative to the thermal state $e^{-\pi|\mathbf{v}|^2}$, satisfies

$$S(f(\cdot, t)) \le \left[\frac{M}{N+M} + \frac{N}{N+M} e^{-t\mu_\rho 2(N+M)/(N+M-1)} \right] S(f_0),$$

[1] Although $\mathcal{L}_{\text{cl}}$ is the generator of the standard Kac model, we look at the evolution it generates on $L^1(\mathbb{R}^{M+N})$ instead of $L^1(\mathbb{S}^{N+M-1}(\sqrt{N+M}))$, see Corollary 1 below.

where

$$\mu_\rho = \int_{-\pi}^{\pi} \rho(\theta)\,\mathrm{d}\theta\ \sin^2(\theta)$$

and ρ is a probability distribution such that (7) *holds.*

One way to understand approach to equilibrium is through entropy production inequalities, i.e., a bound on the ratio of the entropy dissipation and the entropy. This was the approach taken in [29]. These methods do not work in our context. The main reason is that the evolution of the function $f(\mathbf{v}, t)$ is *not* given by a semi group. Therefore, our path for proving Theorem 1 is rather different. First we write the evolution in terms of 'collision histories', i.e., as an average over sequences of collisions. Along each of these we estimate how the entropy decreases. The tool to achieve this is a form of Nelson's hypercontractive estimate. The resulting expressions measure the buildup of correlations between the reservoir and the system. These correlations, however, can be estimated using a sharp version of the Brascamp–Lieb inequalities due to Barthe, see [3].

One can extend the results to three dimensional momentum preserving collisions, however, so far only for a caricature of Maxwellian molecules. Applying this method in the hard sphere and true Maxwellian molecules cases is still unresolved and poses an interesting open problem.

The plan of the paper is as follows: In Section 3 we derive a representation formula for the Kac evolution $\mathrm{e}^{\mathcal{L}t}$ as a sum over collision histories (Theorem 2). The terms in this sum are reminiscent of the Ornstein-Uhlenbeck semi-group. This allows us to prove an entropy inequality based upon Nelson's hypercontractive estimate in Section 4. The resulting expressions estimate the buildup of correlations between the reservoir and the system (Theorem 3). In Section 5 we show how Barthe's sharp version of the geometric Brascamp–Lieb inequality leads to a correlation inequality for the entropy involving marginals, which in turn proves our main entropy inequality. The fact that our Brascamp–Lieb datum is geometric relies on a sum rule which will be proved in Section 6. A short proof of the geometric form of the Brascamp–Lieb inequalities is deferred to Appendix A, as well as some technical details to ensure its applicability in Appendix B. In Section 7 we show how our method can be applied to three-dimensional Maxwellian collisions with a very simple angular dependence.

3. The Representation Formula

The aim of this section is to rewrite (8), that is $e^{\mathcal{L}t}F_0$, in a way which is reminiscent of the Ornstein-Uhlenbeck process. This representation will naturally lead to the next step in the proof of Theorem 1, namely the entropy inequality that will be presented in Theorem 3.

It is convenient to write

$$\mathcal{L} = \Lambda(Q - I),\ \text{where}\ \Lambda = \lambda_S \frac{M}{2} + \lambda_R \frac{N}{2} + \mu M,$$

and the operator Q is a convex combination of R_{ij}s, given by

$$Q = \frac{\lambda_S}{\Lambda(M-1)} \sum_{1\le i<j\le M} R_{ij} + \frac{\lambda_R}{\Lambda(N-1)} \sum_{M<i<j\le N+M} R_{ij} + \frac{\mu}{\Lambda N} \sum_{i=1}^{M} \sum_{j=M+1}^{M+N} R_{ij},$$

i.e., Q is an average over rotation operators. The right hand side of (8) can be written as

$$(e^{\mathcal{L}t} F_0)(\mathbf{v}, \mathbf{w}) = e^{-\Lambda t} \sum_{k=0}^{\infty} \frac{t^k \Lambda^k}{k!} Q^k F_0(\mathbf{v}, \mathbf{w}), \tag{11}$$

where

$$Q^k F_0(\mathbf{v}, \mathbf{w}) = \sum_{\alpha_1, \ldots, \alpha_k} \lambda_{\alpha_1} \cdots \lambda_{\alpha_k} \times \int_{[-\pi,\pi]^k} \rho(\theta_1)\, \mathrm{d}\theta_1 \cdots \rho(\theta_k) \mathrm{d}\theta_k\, F_0 \left(\left[\prod_{l=1}^{k} r_{\alpha_l}(\theta_l) \right]^{-1} (\mathbf{v}, \mathbf{w}) \right). \tag{12}$$

Here, α labels pairs of particles, that is, $\alpha = (i, j)$, $1 \le i < j \le M + N$, $r_\alpha(\theta)$ is defined in (5) and λ_α is given by the rotation corresponding to the index α, that is,

$$\lambda_{(i,j)} = \frac{\lambda_S}{\Lambda(M-1)} \;\text{ if }\; 1 \le i < j \le M,$$
$$\lambda_{(i,j)} = \frac{\lambda_R}{\Lambda(N-1)} \;\text{ if }\; M+1 \le i < j \le M+N,$$
$$\lambda_{(i,j)} = \frac{\mu}{\Lambda N} \;\text{ if }\; 1 \le i \le M,\, M+1 \le j \le M+N\, .$$

Note that the sum over *all* pairs $\sum_\alpha \lambda_\alpha = 1$.

For our purpose, it is convenient to write the function f_0, introduced in (9), as $f_0(\mathbf{v}) = h_0(\mathbf{v}) e^{-\pi|\mathbf{v}|^2}$. Since the Gaussian function is invariant under rotations, (11) takes the form

$$(e^{\mathcal{L}t} F_0)(\mathbf{v}, \mathbf{w}) = e^{-\pi\left(|\mathbf{v}|^2 + |\mathbf{w}|^2\right)} e^{-\Lambda t} \sum_{k=0}^{\infty} \frac{t^k \Lambda^k}{k!} Q^k \left(h_0 \circ P\right)(\mathbf{v}, \mathbf{w}).$$

We introduce the projection $P : \mathbb{R}^{N+M} \to \mathbb{R}^M$ by $P(\mathbf{v}, \mathbf{w}) = \mathbf{v}$, as a reminder that the semigroup $e^{\mathcal{L}t}$ acts on functions that depend on $\mathbf{v}$ as well as $\mathbf{w}$. If we write

$$f(\mathbf{v}, t) = e^{-\pi|\mathbf{v}|^2} h(\mathbf{v}, t),$$

then (10) can be written as

$$h(\mathbf{v}, t) = e^{-\Lambda t} \sum_{k=0}^{\infty} \frac{t^k \Lambda^k}{k!} h_k(\mathbf{v}),$$

where the functions h_k are given by

$$h_k(\mathbf{v}) := \int_{\mathbb{R}^N} Q^k \left(h_0 \circ P\right)(\mathbf{v}, \mathbf{w}) e^{-\pi|\mathbf{w}|^2}\, \mathrm{d}\mathbf{w}.$$

Likewise, the entropy of f is expressed as

$$S(f(\cdot, t)) = \int_{\mathbb{R}^M} h(\mathbf{v}, t) \log h(\mathbf{v}, t) e^{-\pi|\mathbf{v}|^2}\, \mathrm{d}\mathbf{v} =: \mathcal{S}(h(\cdot, t)).$$

Expanding the function $Q^k(h_0 \circ P)(\mathbf{v}, \mathbf{w})$, we find that

$$h_k(\mathbf{v}) = \sum_{\alpha_1,\dots,\alpha_k} \lambda_{\alpha_1} \cdots \lambda_{\alpha_k} \int_{[-\pi,\pi]^k} \rho(\theta_1)\, \mathrm{d}\theta_1 \cdots \rho(\theta_k) \mathrm{d}\theta_k \times \int_{\mathbb{R}^N} (h_0 \circ P) \left(\left[\prod_{l=1}^{k} r_{\alpha_l}(\theta_l) \right]^{-1} (\mathbf{v}, \mathbf{w}) \right) e^{-\pi|\mathbf{w}|^2}\, \mathrm{d}\mathbf{w}, \tag{13}$$

where, as before, see (12), $r_\alpha(\theta)$ rotates the plane given by the index pair α by an angle θ while keeping the other directions fixed. Since $P(\mathbf{v}, \mathbf{w}) = \mathbf{v}$, it is natural to write

$$\left[\prod_{j=1}^{k} r_{\alpha_j}(\theta_j) \right]^{-1} = \begin{pmatrix} A_k(\underline{\alpha}, \underline{\theta}) & B_k(\underline{\alpha}, \underline{\theta}) \\ C_k(\underline{\alpha}, \underline{\theta}) & D_k(\underline{\alpha}, \underline{\theta}) \end{pmatrix},$$

where $A_k \in \mathbb{R}^{M\times M}$ is an $M \times M$ matrix, $B_k \in \mathbb{R}^{M\times N}$, $C_k \in \mathbb{R}^{N\times M}$ and $D_k \in \mathbb{R}^{N\times N}$. Further, $\underline{\alpha} = (\alpha_1, \dots, \alpha_k)$ and $\underline{\theta} = (\theta_1, \dots, \theta_k)$. This notation allows us to rewrite (13) as

$$h_k(\mathbf{v}) = \sum_{\alpha_1,\dots,\alpha_k} \lambda_{\alpha_1} \cdots \lambda_{\alpha_k} \int_{[-\pi,\pi]^k} \rho(\theta_1)\, \mathrm{d}\theta_1 \cdots \rho(\theta_k)\, \mathrm{d}\theta_k \times \int_{\mathbb{R}^N} h_0 \left(A_k(\underline{\alpha}, \underline{\theta})\mathbf{v} + B_k(\underline{\alpha}, \underline{\theta})\mathbf{w} \right) e^{-\pi|\mathbf{w}|^2}\, \mathrm{d}\mathbf{w}.$$

Note that, by the definition of rotations,

$$A_k(\underline{\alpha}, \underline{\theta}) A_k^T(\underline{\alpha}, \underline{\theta}) + B_k(\underline{\alpha}, \underline{\theta}) B_k^T(\underline{\alpha}, \underline{\theta}) = I_M. \tag{14}$$

Lemma 1. *Let $A \in \mathbb{R}^{M\times M}$ and $B \in \mathbb{R}^{M\times N}$ be matrices that satisfy $AA^T + BB^T = I_M$. Then*

$$\int_{\mathbb{R}^N} h(A\mathbf{v} + B\mathbf{w}) e^{-\pi|\mathbf{w}|^2}\, \mathrm{d}\mathbf{w} = \int_{\mathbb{R}^M} h\left(A\mathbf{v} + (I_M - AA^T)^{1/2} \mathbf{u} \right) e^{-\pi|\mathbf{u}|^2}\, \mathrm{d}\mathbf{u}$$

for any integrable function h.

Proof. Denote the range of B by $H \subset \mathbb{R}^M$ and its kernel by $K \subset \mathbb{R}^N$. We may write

$$\int_{\mathbb{R}^N} h(A\mathbf{v} + B\mathbf{w}) e^{-\pi|\mathbf{w}|^2}\, \mathrm{d}\mathbf{w} = \int_K \int_{K^\perp} h(A\mathbf{v} + B\mathbf{u}) e^{-\pi|\mathbf{u}|^2} e^{-\pi|\mathbf{u}'|^2}\, \mathrm{d}\mathbf{u}\mathrm{d}\mathbf{u}' = \int_{K^\perp} h(A\mathbf{v} + B\mathbf{u}) e^{-\pi|\mathbf{u}|^2}\, \mathrm{d}\mathbf{u}.$$

The symmetric map $BB^T : \mathbb{R}^M \to \mathbb{R}^M$ has H as its range and $H^\perp$, that is the orthogonal complement of H in $\mathbb{R}^M$, as its kernel. Indeed, if $BB^T x = 0$ for $x \in \mathbb{R}^M$, then $B^T x = 0$, i.e., $x \in \mathrm{Ker} B^T$ or x is perpendicular to H. This shows that $\mathrm{Ker} BB^T \subset H$. Because $\mathrm{Ran} BB^T \subset H$, if $x \in H^\perp$ then $x \in \left(\mathrm{Ran} BB^T\right)^\perp$ or $x \in \mathrm{Ker} BB^T$. Thus $\mathrm{Ker} BB^T = H^\perp$ and, by a simple dimensional argument, $\mathrm{Ran} BB^T = H$. Hence, the map $BB^T : H \to H$ is invertible (and positive definite). Define the linear map $R : \mathbb{R}^N \to H$ by

$$R = \left(BB^T \right)^{-1/2} B$$

and note that $RR^T = I_H$ while $R^T R$ projects the space $K^\perp$ orthogonally onto H, that is $R^T RK^\perp = H$ while $(R^T R)^2 = R^T R$. Since $K^\perp$ and H have the same dimension, it follows that R^T restricted to H defines an isometry between H and $K^\perp$. Hence,

$$\begin{aligned}\int_{K^\perp} h\,(A\mathbf{v} + B\mathbf{u})\, e^{-\pi|\mathbf{u}|^2}\,\mathrm{d}\mathbf{u} &= \int_{K^\perp} h\left(A\mathbf{v} + \left(BB^T\right)^{1/2} R\mathbf{u}\right) e^{-\pi|\mathbf{u}|^2}\,\mathrm{d}\mathbf{u}\\ &= \int_H h\left(A\mathbf{v} + \left(BB^T\right)^{1/2} RR^T\mathbf{u}\right) e^{-\pi|R^T\mathbf{u}|^2}\,\mathrm{d}\mathbf{u}\\ &= \int_H h\left(A\mathbf{v} + \left(BB^T\right)^{1/2} \mathbf{u}\right) e^{-\pi|\mathbf{u}|^2}\,\mathrm{d}\mathbf{u}.\end{aligned}$$

The assumption $AA^T + BB^T = I_M$, together with the fact that

$$\begin{aligned}&\int_H h\left(A\mathbf{v} + \left(BB^T\right)^{1/2}\mathbf{u}\right) e^{-\pi|\mathbf{u}|^2}\,\mathrm{d}\mathbf{u}\\ &\quad= \int_{H^\perp}\int_H h\left(A\mathbf{v} + (BB^T)^{1/2}\mathbf{u}\right) e^{-\pi|\mathbf{u}|^2}\,\mathrm{d}\mathbf{u}\, e^{-\pi|\mathbf{u}'|^2}\,\mathrm{d}\mathbf{u}',\end{aligned}$$

now implies the lemma. □

The matrix $A_k(\underline{\alpha}, \underline{\theta})$ has an orthogonal singular value decomposition,

$$A_k(\underline{\alpha}, \underline{\theta}) = U_k(\underline{\alpha}, \underline{\theta})\Gamma_k(\underline{\alpha}, \underline{\theta})V_k^T(\underline{\alpha}, \underline{\theta}), \tag{15}$$

where $\Gamma_k(\underline{\alpha}, \underline{\theta}) = \operatorname{diag}[\gamma_{k,1}(\underline{\alpha}, \underline{\theta}), \ldots, \gamma_{k,M}(\underline{\alpha}, \underline{\theta})]$ is the diagonal matrix whose entries $\gamma_{k,j}(\underline{\alpha}, \underline{\theta})$, $j = 1, \ldots, M$, are the singular values of $A_k(\underline{\alpha}, \underline{\theta})$, and $U_k(\underline{\alpha}, \underline{\theta})$ and $V_k(\underline{\alpha}, \underline{\theta})$ are orthogonal $M \times M$ matrices. Note that (14) implies $\gamma_{k,j}(\underline{\alpha}, \underline{\theta}) \in [0, 1]$ for $j = 1, \ldots, M$. We shall use the abbreviation

$$h_0(U_k(\underline{\alpha}, \underline{\theta})\mathbf{v}) = h_{0,U_k(\underline{\alpha},\underline{\theta})}(\mathbf{v}).$$

These considerations can be summarized by the representation formula presented in the following theorem.

Theorem 2 (Representation formula). *The function h_k can be written as*

$$\begin{aligned}h_k(\mathbf{v}) = \sum_{\alpha_1,\ldots,\alpha_k} \lambda_{\alpha_1}\cdots\lambda_{\alpha_k} \int_{[-\pi,\pi]^k} \rho(\theta_1)\,\mathrm{d}\theta_1\cdots\rho(\theta_k)\,\mathrm{d}\theta_k \times\\ \times \int_{\mathbb{R}^M} h_{0,U_k(\underline{\alpha},\underline{\theta})}\left(\Gamma_k(\underline{\alpha}, \underline{\theta})V_k^T(\underline{\alpha}, \underline{\theta})\mathbf{v} + \left(I_M - \Gamma_k^2(\underline{\alpha}, \underline{\theta})\right)^{1/2}\mathbf{w}\right) e^{-\pi|\mathbf{w}|^2}\,\mathrm{d}\mathbf{w},\end{aligned} \tag{16}$$

where $h_{0,U_k(\underline{\alpha},\underline{\theta})}$, $\Gamma_k(\underline{\alpha}, \underline{\theta})$ and V_k are as defined above.

4. The Hypercontractive Estimate

Starting from (16) and using convexity of the entropy and Jensen's inequality together with

$$\sum_{\alpha_1,\ldots,\alpha_k} \lambda_{\alpha_1} \cdots \lambda_{\alpha_k} \int_{[-\pi,\pi]^k} \rho(\theta_1)\,\mathrm{d}\theta_1 \cdots \rho(\theta_k)\,\mathrm{d}\theta_k = 1,$$

we get

$$\mathcal{S}(h_k) \le \sum_{\alpha_1,\ldots,\alpha_k} \lambda_{\alpha_1} \cdots \lambda_{\alpha_k} \int_{[-\pi,\pi]^k} \rho(\theta_1)\,\mathrm{d}\theta_1 \cdots \rho(\theta_k)\,\mathrm{d}\theta_k\, \mathcal{S}(g_k(\cdot, \underline{\alpha}, \underline{\theta})),$$

where we set

$$g_k(\mathbf{v}, \underline{\alpha}, \underline{\theta}) = \int_{\mathbb{R}^M} h_{0,U_k(\underline{\alpha},\underline{\theta})} \left(\Gamma_k(\underline{\alpha}, \underline{\theta})\mathbf{v} + \left(I_M - \Gamma_k^2(\underline{\alpha}, \underline{\theta}) \right)^{1/2} \mathbf{w} \right) e^{-\pi|\mathbf{w}|^2}\,\mathrm{d}\mathbf{w}, \tag{17}$$

and we removed the rotation $V_k^T(\underline{\alpha}, \underline{\theta})$ by a change of variables.

To explain the main observation in this section we look at (17) when $M = 1$. Since $0 \le \gamma_k(\underline{\alpha}, \underline{\theta}) \le 1$, we can write $\gamma_k(\underline{\alpha}, \underline{\theta}) = e^{-t}$ and we get $g_k(v, \underline{\alpha}, \underline{\theta}) = N_t(h_{0,U_k(\underline{\alpha},\underline{\theta})})$ where N_t is the Ornstein-Uhlenbeck semigroup, that is

$$N_t h(x) = \int_{\mathbb{R}} h\left(e^{-t}x + \sqrt{1-e^{-2t}}\,y\right) e^{-\pi y^2}\,\mathrm{d}y.$$

Thus Theorem 2 renders the function h_k as a convex combination of terms reminiscent of the Ornstein-Uhlenbeck process, albeit in matrix form. We make use of this observation to find a bound for $\mathcal{S}(g_k(\cdot, \underline{\alpha}, \underline{\theta}))$. This bound together with a suitable correlation inequality proved in the next section will lead to a bound for $\mathcal{S}(h_k)$.

In addition to the notation developed in the previous section, we need various marginals of the function $h_{0,U_k(\underline{\alpha},\underline{\theta})}$. Quite generally, if h is a function of M variables and $\sigma \subset \{1, \ldots, M\}$, we shall denote by h^σ the marginals of h with respect to the variables v_j with weight $e^{-\pi v_j^2}$, $j \in \sigma$. For instance, we have

$$h^{\{1,2\}}(v_3, \ldots, v_M) = \int_{\mathbb{R}^2} h(v_1, v_2, v_3, \ldots, v_M) e^{-\pi\left(v_1^2+v_2^2\right)}\,\mathrm{d}v_1\mathrm{d}v_2,$$

and for $\sigma = \emptyset$ we set $h^\sigma = h$. It will be convenient to use the matrix $P_\sigma : \mathbb{R}^M \to \mathbb{R}^{|\sigma|}$ that projects $\mathbb{R}^M$ orthogonaly onto $\mathbb{R}^{|\sigma|}$ which we will identify with subspace of $\mathbb{R}^M$. To give an example, let $\mathbf{v} = (v_1, ..., v_M)$. Then $P_{\{1,2\}}\mathbf{v} = (v_1, v_2)$. The following theorem is the main result of this section.

Theorem 3 (Partial entropy bound). *Let $h_0 \in L^1(\mathbb{R}^M, e^{-\pi|\mathbf{v}|^2}d\mathbf{v})$ be nonnegative and assume that $\mathcal{S}(h_0) < \infty$. Then*

$$\begin{aligned}
&\mathcal{S}(g_k(\cdot, \underline{\alpha}, \underline{\theta})) \\
&\quad \le \sum_{\sigma\subset\{1,\ldots,M\}} \prod_{i\in\sigma^c} \gamma_{k,i}^2 \prod_{j\in\sigma} \left(1 - \gamma_{k,j}^2\right) \\
&\qquad \times \int_{\mathbb{R}^M} h_0(\mathbf{v}) \log h_{0,U_k(\underline{\alpha},\underline{\theta})}^{\sigma} \left(P_{\sigma^c} U_k(\underline{\alpha}, \underline{\theta})^T \mathbf{v}\right) e^{-\pi|\mathbf{v}|^2}\,\mathrm{d}\mathbf{v},
\end{aligned} \tag{18}$$

where σ^c is the complement of the set σ in $\{1, ..., M\}$.

A key role in the proof of Theorem 3 is played by Nelson's hypercontractive estimate.

Theorem 4 (Nelson's hypercontractive estimate). *The Ornstein-Uhlenbeck semigroup,*

$$N_t h(x) = \int_{\mathbb{R}} h\left(e^{-t}x + \sqrt{1-e^{-2t}}\,y\right) e^{-\pi y^2}\,\mathrm{d}y,$$

for $t \geq 0$, *is bounded from* $L^p(\mathbb{R}, e^{-\pi x^2}\,\mathrm{d}x)$ *to* $L^q(\mathbb{R}, e^{-\pi x^2}\,\mathrm{d}x)$ *if and only if*

$$(p-1) \geq e^{-2t}(q-1).$$

For such values of p *and* q*,*

$$\|N_t h\|_q \leq \|h\|_p$$

with equality if and only if h *is constant.*[2]

Proof. For a proof we refer the reader to [27]. For other proofs see [12,15,17,18]. □

Nelson's hypercontractive estimate, that is Theorem 4, implies the following Corollary, which will be useful in the proof of Theorem 3.

Corollary 2 (Entropic version of Nelson's hypercontractive estimate). *Let* $h : \mathbb{R} \to \mathbb{R}_+$ *be a function in* $L^1(\mathbb{R}, e^{-\pi x^2}\,\mathrm{d}x)$ *with finite entropy, i.e.,*

$$\mathcal{S}(h) = \int_{\mathbb{R}} h(x) \log h(x)\, e^{-\pi x^2} dx < \infty.$$

Then

$$\mathcal{S}(N_t h) \leq e^{-2t}\mathcal{S}(h) + (1-e^{-2t})\|h\|_1 \log \|h\|_1$$

for all $t \geq 0$.

Proof. Let $h \in L^p(\mathbb{R}, e^{-\pi x^2}\,\mathrm{d}x)$, for $p > 1$ small, be a nonnegative function. As $\|N_t h\|_1 = \|h\|_1$, we can apply Nelson's hypercontractive estimate, which implies that for p, q that satisfy $(p-1) = e^{-2t}(q-1)$,

$$\frac{\|N_t h\|_q - \|N_t h\|_1}{q-1} \leq \frac{\|h\|_p - \|h\|_1}{q-1} = e^{-2t}\frac{\|h\|_p - \|h\|_1}{p-1}.$$

Sending $p \to 1$ and hence $q \to 1$, we get the claimed estimate for such functions h. If h just has finite entropy one cuts off h at large values, uses the above estimate and removes the cutoff using the monotone convergence theorem. □

We are now ready to prove Theorem 3.

[2] Here and in the following, if there is no room for confusion, we denote by $\|h\|_p = \|h\|_{L^p(\mathbb{R}, e^{-\pi x^2} dx)}$ the L^p norm of h with respect to the Gaussian measure on R. Note also that $L^1(\mathbb{R}, e^{-\pi x^2}\,\mathrm{d}x) \subset L^p(\mathbb{R}, e^{-\pi x^2}\,\mathrm{d}x)$ for all $p \geq 1$ since $e^{-\pi x^2}\,\mathrm{d}x$ is a probability measure.

Proof of Theorem 3. Remember that $0 \le \gamma_{k,j}(\underline{\alpha}, \underline{\theta}) \le 1$ for $j = 1, ..., M$. Thus, by inductively applying Corollary 2 to

$$\int_{\mathbb{R}^M} h_{0,U_k(\underline{\alpha},\underline{\theta})}\left(\gamma_{k,1}v_1 + \sqrt{1-\gamma_{k,1}^2}\,u_1, \ldots, \gamma_{k,M}v_M + \sqrt{1-\gamma_{k,M}^2}\,u_M\right) \times e^{-\pi\sum_{j=1}^M u_j^2}\,\mathrm{d}u_1\cdots\mathrm{d}u_M,$$

we obtain

$$\mathcal{S}(g_k(\cdot,\underline{\alpha},\underline{\theta})) \le \sum_{\sigma\subset\{1,\ldots,M\}}\prod_{i\in\sigma^c}\gamma_{k,i}^2\prod_{j\in\sigma}(1-\gamma_{k,j}^2)\int_{\mathbb{R}^{|\sigma^c|}} h^\sigma_{0,U_k(\underline{\alpha},\underline{\theta})}(\mathbf{u})\log h^\sigma_{0,U_k(\underline{\alpha},\underline{\theta})}(\mathbf{u})\,e^{-\pi|\mathbf{u}|^2}\,\mathrm{d}\mathbf{u}.$$

Inserting the definition of the marginal $h^\sigma_{0,U_k(\underline{\alpha},\underline{\theta})}$, we see that

$$\begin{aligned}
&\int_{\mathbb{R}^{|\sigma^c|}} h^\sigma_{0,U_k(\underline{\alpha},\underline{\theta})}(\mathbf{u})\log h^\sigma_{0,U_k(\underline{\alpha},\underline{\theta})}(\mathbf{u})\,e^{-\pi|\mathbf{u}|^2}\,\mathrm{d}\mathbf{u}\\
&\quad= \int_{\mathbb{R}^M} h^\sigma_{0,U_k(\underline{\alpha},\underline{\theta})}(P_{\sigma^c}\mathbf{v})\log h^\sigma_{0,U_k(\underline{\alpha},\underline{\theta})}(P_{\sigma^c}\mathbf{v})\,e^{-\pi|\mathbf{v}|^2}\,\mathrm{d}\mathbf{v}\\
&\quad= \int_{\mathbb{R}^M} h_{0,U_k(\underline{\alpha},\underline{\theta})}(\mathbf{v})\log h^\sigma_{0,U_k(\underline{\alpha},\underline{\theta})}(P_{\sigma^c}\mathbf{v})\,e^{-\pi|\mathbf{v}|^2}\,\mathrm{d}\mathbf{v}\\
&\quad= \int_{\mathbb{R}^M} h_0(\mathbf{v})\log h^\sigma_{0,U_k(\underline{\alpha},\underline{\theta})}(P_{\sigma^c}U_k(\underline{\alpha},\underline{\theta})^T\mathbf{v})\,e^{-\pi|\mathbf{v}|^2}\,\mathrm{d}\mathbf{v},
\end{aligned}$$

which finishes the proof of Theorem 3. □

5. The Key Entropy Bound

Collecting the results of the previous sections we get the following bound

$$\begin{aligned}
\mathcal{S}(h_k) \le &\sum_{\alpha_1,\ldots,\alpha_k}\lambda_{\alpha_1}\cdots\lambda_{\alpha_k}\int_{[-\pi,\pi]^k}\rho(\theta_1)\,\mathrm{d}\theta_1\cdots\rho(\theta_k)\,\mathrm{d}\theta_k\\
&\times\sum_{\sigma\subset\{1,\ldots,M\}}\prod_{i\in\sigma^c}\gamma_{k,i}^2\prod_{j\in\sigma}\left(1-\gamma_{k,j}^2\right)\\
&\times\int_{\mathbb{R}^M} h_0(\mathbf{v})\log h^\sigma_{0,U_k(\underline{\alpha},\underline{\theta})}\left(P_{\sigma^c}U_k(\underline{\alpha},\underline{\theta})^T\mathbf{v}\right)\,e^{-\pi|\mathbf{v}|^2}\,\mathrm{d}\mathbf{v}.
\end{aligned}\tag{19}$$

The right-hand side of (19) contains a large sum over the entropy of marginals of h_0. In order to bound such a sum in terms of the entropy of h_0 one may try to apply some version of the Loomis-Whitney inequality [23] or, more precisely, of an inequality by Han [19]. This is essentially correct, but will require a substantial generalization of this inequality. Let us first formulate the main theorem of this section.

Theorem 5 (Entropy bound). *The estimate*

$$\mathcal{S}(h_k) \le \left[\frac{M}{N+M} + \frac{N}{N+M}\left(1-\mu_\rho\frac{N+M}{N\Lambda}\right)^k\right]\mathcal{S}(h_0) \tag{20}$$

holds.

The generalization of Han's inequality mentioned above was proven by Carlen-Cordero-Erausquin in [11]. It is based on the geometric Brascamp–Lieb inequality due to Ball [1], see also [2], in the rank one case, and due to Barthe [3] in the general case.

Theorem 6 (Correlation inequality). *For $i = 1, \dots K$, let $H_i \subset \mathbb{R}^M$ be subspaces of dimension d_i and $B_i : \mathbb{R}^M \to H_i$ be linear maps with the property that $B_i B_i^T = I_{H_i}$, the identity map on H_i. Assume further that there are non-negative constants c_i, $i = 1, \dots, K$ such that*

$$\sum_{i=1}^{K} c_i B_i^T B_i = I_M. \tag{21}$$

Then, for nonnegative functions $f_i : H_i \to \mathbb{R}$,

$$\int_{\mathbb{R}^M} \prod_{i=1}^{K} f_i^{c_i}(B_i \mathbf{v})\, e^{-\pi|\mathbf{v}|^2}\, \mathrm{d}\mathbf{v} \le \prod_{i=1}^{K} \left(\int_{H_i} f_i(u)\, e^{-\pi|u|^2}\, \mathrm{d}u \right)^{c_i}. \tag{22}$$

Moreover,

$$\begin{aligned} &\int_{\mathbb{R}^M} h(\mathbf{v}) \log h(\mathbf{v})\, e^{-\pi|\mathbf{v}|^2}\, \mathrm{d}\mathbf{v} \\ &\ge \sum_{i=1}^{K} c_i \left[\int_{\mathbb{R}^M} h(\mathbf{v}) \log f_i(B_i \mathbf{v})\, e^{-\pi|\mathbf{v}|^2}\, \mathrm{d}\mathbf{v} - \log \int_{H_i} f_i(u)\, e^{-\pi|u|^2}\, \mathrm{d}u \right], \end{aligned} \tag{23}$$

for any nonnegative function $h \in L^1(\mathbb{R}^M, e^{-\pi|\mathbf{v}|^2} d\mathbf{v})$.

Since Theorem 6 is very useful in a number of applications, and for the readers convenience, we will give an elementary proof in Appendix A.

Remark 2. By taking the trace in (21) one sees that

$$\sum_{i=1}^{K} c_i d_i = M.$$

We would like to apply (23) to the right hand side of (19). An immediate problem is that (19) is in terms of integrals and not sums. While there are some results available for continuous indices (see, e.g., [4]), they do not apply to our situation and hence we will take a more direct approach and approximate the measure $\rho(\theta)\mathrm{d}\theta$ by a discrete measure. It is important that the approximation also satisfies the constraint (7). The following lemma establishes such an approximation. Its proof is given in Appendix B.

Lemma 2. *Let ρ be a probability density on $[-\pi, \pi]$ whose Fourier series converges absolutely and assume that (7) is satisfied. There exists a sequence of discrete probability measures ν_K, $K = 1, 2, \dots,$ such that for every continuous function f on $[-\pi, \pi]$*

$$\lim_{K \to \infty} \int_{-\pi}^{\pi} f(\theta)\, \nu_K(\mathrm{d}\theta) = \int_{-\pi}^{\pi} f(\theta) \rho(\theta)\, \mathrm{d}\theta.$$

Moreover,

$$\int_{-\pi}^{\pi} \cos\theta \sin\theta\, \nu_K(\mathrm{d}\theta) = 0,$$

for all $K \in \mathbb{N}$. *More precisely,*

$$\nu_K(\mathrm{d}\theta) = \frac{2\pi}{4K+1} \sum_{\ell=-2K}^{2K} \rho_K\left(\frac{2\pi\ell}{4K+1}\right) \delta\left(\theta - \frac{2\pi\ell}{4K+1}\right) \mathrm{d}\theta,$$

where $\rho_K(\theta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \rho(\theta-\phi)\, p_K(\phi)\,\mathrm{d}\phi$ *and* p_K *is the Fejér kernel*

$$p_K(\theta) := \frac{1}{2K+1}\left(\sum_{k=-K}^{K} e^{ik\theta}\right)^2 = \frac{1}{2K+1}\left(\frac{\sin\left((K+\frac{1}{2})\theta\right)}{\sin\frac{\theta}{2}}\right)^2.$$

At this point we can prepare the ground for the application of Theorem 6 to inequality (19). We first replace $\rho(\theta)\mathrm{d}\theta$ in (19) with $\nu_K(\mathrm{d}\theta)$. Setting

$$\omega_{\ell_j} = \rho_K(\theta_{\ell_j}),\ \theta_{\ell_j} = \frac{2\pi\ell_j}{4K+1},\ \text{and } \underline{\theta}_K = (\theta_{\ell_1},\ldots,\theta_{\ell_k}),$$

we obtain

$$\begin{aligned}
&\sum_{\alpha_1,\ldots,\alpha_k} \lambda_{\alpha_1}\cdots\lambda_{\alpha_k} \int_{[-\pi,\pi]^k} \nu_K(\mathrm{d}\theta_1)\cdots\nu_K(\mathrm{d}\theta_k) \sum_{\sigma\subset\{1,\ldots,M\}} \prod_{i\in\sigma^{\mathrm{c}}} \gamma_{k,i}(\underline{\alpha},\underline{\theta})^2 \\
&\quad\times \prod_{j\in\sigma}\left(1-\gamma_{k,j}(\underline{\alpha},\underline{\theta})^2\right) \int_{\mathbb{R}^M} h_0(\mathbf{v}) \log h^{\sigma}_{0,U_k(\underline{\alpha},\underline{\theta})}(P_{\sigma^{\mathrm{c}}} U_k(\underline{\alpha},\underline{\theta})^T \mathbf{v})\, e^{-\pi|\mathbf{v}|^2}\,\mathrm{d}\mathbf{v} \\
&= \sum_{\alpha_1,\ldots,\alpha_k} \lambda_{\alpha_1}\cdots\lambda_{\alpha_k} \sum_{-2K\le\ell_1,\ldots,\ell_k\le 2K} \prod_{j=1}^{k}\omega_{\ell_j} \sum_{\sigma\subset\{1,\ldots,M\}} \prod_{i\in\sigma^{\mathrm{c}}} \gamma_{k,i}(\underline{\alpha},\underline{\theta}_K)^2 \\
&\quad\times \prod_{j\in\sigma}\left(1-\gamma_{k,j}(\underline{\alpha},\underline{\theta}_K)^2\right) \int_{\mathbb{R}^M} h_0(\mathbf{v}) \log h^{\sigma}_{0,U_k(\underline{\alpha},\underline{\theta}_K)}(P_{\sigma^{\mathrm{c}}} U_k(\underline{\alpha},\underline{\theta}_K)^T \mathbf{v})\, e^{-\pi|\mathbf{v}|^2}\,\mathrm{d}\mathbf{v}.
\end{aligned} \tag{24}$$

In order to apply Theorem 6 to (24) we have to replace the sum over the index i with a sum over the indices $\alpha_1,\ldots,\alpha_k,\ell_1,\ldots\ell_k$ and all subsets $\sigma\subset\{1,\ldots,M\}$. Moreover, we substitute

the constants c_i	by $\frac{1}{C_{k,M}}\lambda_{\alpha_1}\cdots\lambda_{\alpha_k}\prod_{j=1}^{k}\omega_{\ell_j}\prod_{i\in\sigma^{\mathrm{c}}}\gamma_{k,i}(\underline{\alpha},\underline{\theta}_K)^2\prod_{j\in\sigma}(1-\gamma_{k,j}(\underline{\alpha},\underline{\theta}_K)^2)$, where $C_{k,M}$ will be defined in Theorem 7 below,
the functions $f_i(\mathbf{w})$	by $h^{\sigma}_{0,U_k(\underline{\alpha},\underline{\theta}_K)}(\mathbf{w})$,
the linear maps B_i	by $P_{\sigma^{\mathrm{c}}}U_k(\underline{\alpha},\underline{\theta}_K)^T$,
the functions $f_i(B_i\mathbf{v})$	by $h^{\sigma}_{0,U_k(\underline{\alpha},\underline{\theta}_K)}(P_{\sigma^{\mathrm{c}}}U_k(\underline{\alpha},\underline{\theta}_K)^T\mathbf{v})$,
and the subspaces H_i	by $\mathbb{R}^{\lvert\sigma^c\rvert}$.

For any given index i the condition $B_i B_i^T = I_{H_i}$ corresponds to

$$P_{\sigma^c} U_k(\underline{\alpha}, \underline{\theta}_K)^T U_k(\underline{\alpha}, \underline{\theta}_K) P_{\sigma^c} \;=\; P_{\sigma^c}$$

which is the identity on $\mathbb{R}^{|\sigma^c|}$.

The next theorem establishes the sum rule (21) in our setting and hence ensures the applicability of Theorem 6 to (24).

Theorem 7 (The sum rule). *If $\nu(\mathrm{d}\theta)$ is a probability measure satisfying (7), then*

$$\begin{aligned}
&\sum_{\alpha_1,\dots,\alpha_k} \lambda_{\alpha_1} \cdots \lambda_{\alpha_k} \int_{[-\pi,\pi]^k} \nu(\mathrm{d}\theta_1) \cdots \nu(\mathrm{d}\theta_k) \\
&\quad \times \sum_{\sigma \subset \{1,\dots,M\}} \prod_{i \in \sigma^c} \gamma_{k,i}(\underline{\alpha}, \underline{\theta})^2 \prod_{j \in \sigma} \left(1 - \gamma_{k,j}(\underline{\alpha}, \underline{\theta})^2\right) U_k(\underline{\alpha}, \underline{\theta}) P_{\sigma^c}^T P_{\sigma^c} U_k(\underline{\alpha}, \underline{\theta})^T \\
&\quad = C_{k,M} I_M,
\end{aligned} \tag{25}$$

where

$$C_{k,M} = \left[\frac{M}{N+M} + \frac{N}{N+M} \left(1 - \mu_\nu \frac{N+M}{N\Lambda}\right)^k \right]$$

with

$$\mu_\nu = \mu \int \nu(\mathrm{d}\theta) \sin^2\theta.$$

The proof will be given in Section 6. We observe here that it follows from Theorem 2 that $\mu_\rho = \lim_{K\to\infty} \mu_{\nu_K}$.

Proof of Theorem 5. First, we consider the case where ρ is repaced by ν_K and use Theorem 6 together with Theorem 7 and the identification rules described above. The entropy inequality (23) now says that

$$\begin{aligned}
&\int_{\mathbb{R}^M} h_0(\mathbf{v}) \log h_0(\mathbf{v}) e^{-\pi|\mathbf{v}|^2} \,\mathrm{d}\mathbf{v} \\
&\geq \frac{1}{C_{k,M}} \sum_{\alpha_1,\dots,\alpha_k} \lambda_{\alpha_1} \cdots \lambda_{\alpha_k} \sum_{-2K \le \ell_1,\dots,\ell_k \le 2K} \prod_{j=1}^k \omega_{\ell_j} \sum_{\sigma \subset \{1,\dots,M\}} \prod_{i\in\sigma^c} \gamma_{k,i}(\underline{\alpha}, \underline{\theta}_K)^2 \\
&\quad \times \prod_{j\in\sigma} \left(1 - \gamma_{k,j}(\underline{\alpha}, \underline{\theta}_K)^2\right) \Bigg[\int_{\mathbb{R}^M} h_0(\mathbf{v}) \log h^\sigma_{0,U_k(\underline{\alpha},\underline{\theta}_K)}(P_{\sigma^c} U_k(\underline{\alpha}, \underline{\theta}_K)^T \mathbf{v})\, e^{-\pi|\mathbf{v}|^2} \,\mathrm{d}\mathbf{v} \\
&\qquad\qquad - \log \int_{\mathbb{R}^{|\sigma^c|}} h^\sigma_{0,U_k(\underline{\alpha},\underline{\theta}_K)}(u)\, e^{-\pi|u|^2} \,\mathrm{d}u \Bigg].
\end{aligned}$$

However, since h_0 is normalized and $U_k(\underline{\alpha}, \underline{\theta}_K)$ is orthogonal, we find that

$$\begin{aligned}
\int_{\mathbb{R}^{|\sigma^c|}} h^\sigma_{0,U_k(\underline{\alpha},\underline{\theta}_K)}(u)\, e^{-\pi|u|^2} \,\mathrm{d}\,u &= \int_{\mathbb{R}^{|\sigma^c|}} \int_{\mathbb{R}^{|\sigma|}} h_{0,U_k(\underline{\alpha},\underline{\theta}_K)}(v, u)\, e^{-\pi|v|^2} \,\mathrm{d}v\, e^{-\pi|u|^2} \,\mathrm{d}u \\
&= \int_{\mathbb{R}^M} h_0(U_k(\underline{\alpha}, \underline{\theta}_K)\mathbf{v})\, e^{-\pi|\mathbf{v}|^2} \,\mathrm{d}\mathbf{v} \\
&= 1.
\end{aligned}$$

Thus, we find

$$\sum_{\alpha_1,\ldots,\alpha_k} \lambda_{\alpha_1} \cdots \lambda_{\alpha_k} \sum_{-2K \le \ell_1,\ldots,\ell_k \le 2K} \prod_{j=1}^{k} \omega_{\ell_j}$$
$$\times \sum_{\sigma \subset \{1,\ldots,M\}} \prod_{i \in \sigma^c} \gamma_{k,i}(\underline{\alpha}, \underline{\theta}_K)^2 \prod_{j \in \sigma} \left(1 - \gamma_{k,j}(\underline{\alpha}, \underline{\theta}_K)^2\right)$$
$$\times \int_{\mathbb{R}^M} h_0(\mathbf{v}) \log h^{\sigma}_{0,U_k(\underline{\alpha},\underline{\theta}_K)}(P_{\sigma^c} U_k(\underline{\alpha}, \underline{\theta}_K)^T \mathbf{v})\, e^{-\pi|\mathbf{v}|^2}\, \mathrm{d}\mathbf{v}$$
$$\le C_{k,M} \mathcal{S}(h_0)\,. \tag{26}$$

As $K \to \infty$, the left-hand side of (26) converges to the right-hand side of (19). □

We now have all ingredients to give the proof of Theorem 1.

Proof of Theorem 1. Recall from Section 3, that

$$f(\mathbf{v}, t) = e^{-\pi|\mathbf{v}|^2} e^{-\Lambda t} \sum_{k=0}^{\infty} \frac{t^k \Lambda^k}{k!} h_k(\mathbf{v}),$$

and that $S(f(\cdot, t)) = \mathcal{S}(h(\cdot, t))$. Combining Theorem 3 and Theorem 5, we obtain

$$\mathcal{S}(h_k) \le C_{k,M} \mathcal{S}(h_0),$$

by convexity of the entropy, and computing

$$e^{-\Lambda t} \sum_{k=0}^{\infty} \frac{\Lambda^k t^k}{k!} C_{k,M}$$

yields Theorem 1. □

6. The Sum Rule. Proof of Theorem 7

We have to compute the matrix

$$Z := \sum_{\alpha_1,\ldots,\alpha_k} \lambda_{\alpha_1} \cdots \lambda_{\alpha_k} \int_{[-\pi,\pi]^k} \nu(\mathrm{d}\theta_1) \cdots \nu(\mathrm{d}\theta_k)$$
$$\times \sum_{\sigma \subset \{1,\ldots,M\}} \prod_{i \in \sigma^c} \gamma_{k,i}(\underline{\alpha}, \underline{\theta})^2 \prod_{j \in \sigma} \left(1 - \gamma_{k,j}(\underline{\alpha}, \underline{\theta})^2\right) U_k(\underline{\alpha}, \underline{\theta}) P^T_{\sigma^c} P_{\sigma^c} U_k(\underline{\alpha}, \underline{\theta})^T.$$

Obviously $P^T_{\sigma^c} P_{\sigma^c} = P_{\sigma^c}$ and hence

$$Z = \sum_{\alpha_1,\ldots,\alpha_k} \lambda_{\alpha_1} \cdots \lambda_{\alpha_k} \int_{[-\pi,\pi]^k} \nu(\mathrm{d}\theta_1) \cdots \nu(\mathrm{d}\theta_k)$$
$$\times U_k(\underline{\alpha}, \underline{\theta}) \left[\sum_{\sigma \subset \{1,\ldots,M\}} \prod_{i \in \sigma^c} \gamma_{k,i}(\underline{\alpha}, \underline{\theta})^2 \prod_{j \in \sigma} \left(1 - \gamma_{k,j}(\underline{\alpha}, \underline{\theta})^2\right) P_{\sigma^c} \right] U_k(\underline{\alpha}, \underline{\theta})^T.$$

The sum on σ is easily evaluated and yields the matrix $\Gamma_k^2(\underline{\alpha}, \underline{\theta})$. Hence, recalling the orthogonal singular value decomposition (15) of $A_k(\underline{\alpha}, \underline{\theta})$, that is, $A_k(\underline{\alpha}, \underline{\theta}) = U_k(\underline{\alpha}, \underline{\theta})\Gamma_k(\underline{\alpha}, \underline{\theta})V_k^T(\underline{\alpha}, \underline{\theta})$, we find that

$$Z = \sum_{\alpha_1,\ldots,\alpha_k} \lambda_{\alpha_1} \cdots \lambda_{\alpha_k} \int_{[-\pi,\pi]^k} \nu(\mathrm{d}\theta_1) \cdots \nu(\mathrm{d}\theta_k)\, A_k(\underline{\alpha}, \underline{\theta}) A_k^T(\underline{\alpha}, \underline{\theta}). \tag{27}$$

One can think about this expression in the following fashion. Recall that

$$\left[\prod_{l=1}^k r_{\alpha_l}(\theta_l)\right]^{-1} = \begin{pmatrix} A_k(\underline{\alpha}, \underline{\theta}) & B_k(\underline{\alpha}, \underline{\theta}) \\ C_k(\underline{\alpha}, \underline{\theta}) & D_k(\underline{\alpha}, \underline{\theta}) \end{pmatrix}.$$

With this notation, the matrix Z equals the top left entry of the matrix

$$\sum_{\alpha_1,\ldots,\alpha_k} \lambda_{\alpha_1} \cdots \lambda_{\alpha_k} \int_{[-\pi,\pi]^k} \nu(\mathrm{d}\theta_1) \cdots \nu(\mathrm{d}\theta_k) \left[\prod_{l=1}^k r_{\alpha_l}(\theta_l)\right]^{-1} \begin{pmatrix} I_M & 0 \\ 0 & 0 \end{pmatrix} \left[\prod_{l=1}^k r_{\alpha_l}(\theta_l)\right].$$

The computation hinges on a repeated application of the elementary identity

$$\begin{aligned} &\int_{-\pi}^{\pi} \nu(\mathrm{d}\theta) \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} m_1 & 0 \\ 0 & m_2 \end{pmatrix} \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} \\ &= \begin{pmatrix} (1-\tilde{\nu})m_1 + \tilde{\nu}m_2 & 0 \\ 0 & (1-\tilde{\nu})m_2 + \tilde{\nu}m_1 \end{pmatrix}, \end{aligned}$$

where $\tilde{\nu} = \int \nu(\mathrm{d}\theta)\, \sin^2(\theta)$. For this to be true we just need (7). We easily check that for the rotations $r_\alpha(\theta)$

$$\begin{aligned} &\sum_\alpha \lambda_\alpha \int_{-\pi}^{\pi} \nu(\mathrm{d}\theta)\, r_\alpha(\theta)^{-1} \begin{pmatrix} m_1 I_M & 0 \\ 0 & m_2 I_N \end{pmatrix} r_\alpha(\theta) \\ &= \frac{1}{\Lambda}\left(\frac{M\lambda_S}{2} + \frac{N\lambda_R}{2}\right) \begin{pmatrix} m_1 I_M & 0 \\ 0 & m_2 I_N \end{pmatrix} \\ &\quad + \frac{\mu}{\Lambda N} \begin{pmatrix} N(M-1)m_1 + N((1-\tilde{\nu})m_1 + \tilde{\nu}m_2) I_M & 0 \\ 0 & (N-1)Mm_2 + M(\tilde{\nu}m_1 + (1-\tilde{\nu})m_2) I_N \end{pmatrix} \\ &= \begin{pmatrix} m_1 I_M & 0 \\ 0 & m_2 I_N \end{pmatrix} + \frac{\mu_\nu}{\Lambda N} \begin{pmatrix} N(m_2 - m_1) I_M & 0 \\ 0 & M(m_1 - m_2) I_N \end{pmatrix} \end{aligned} \tag{28}$$

where $\mu_\nu = \tilde{\nu}\mu$. Denote by $L(\nu_1, \nu_2)$ the $(N+M) \times (N+M)$ matrix

$$L(m_1, m_2) = \begin{pmatrix} m_1 I_M & 0 \\ 0 & m_2 I_N \end{pmatrix},$$

and set

$$\mathcal{P} = I_2 - \frac{\mu_\nu}{\Lambda N} \begin{pmatrix} N & -N \\ -M & M \end{pmatrix}.$$

Then (28) is recast as

$$\sum_\alpha \lambda_\alpha \int_{-\pi}^{\pi} \nu(\mathrm{d}\theta)\, r_\alpha(\theta)^{-1} L(m_1, m_2) r_\alpha(\theta) = L(m_1', m_2'), \tag{29}$$

where

$$\begin{pmatrix} m_1' \\ m_2' \end{pmatrix} = \mathcal{P} \begin{pmatrix} m_1 \\ m_2 \end{pmatrix}.$$

By a repeated application of (29) we obtain

$$\sum_{\alpha_1,\ldots,\alpha_k} \lambda_{\alpha_1} \cdots \lambda_{\alpha_k} \int_{[-\pi,\pi]^k} \nu(\mathrm{d}\theta_1) \cdots \nu(\mathrm{d}\theta_k) \left[\prod_{j=1}^k r_{\alpha_j}(\theta_j)\right]^T L(\underline{m}) \left[\prod_{j=1}^k r_{\alpha_j}(\theta_j)\right]$$
$$= L(\mathcal{P}^k \underline{m}).$$

Thus,

$$Z = \left(\mathcal{P}^k \begin{pmatrix} 1 \\ 0 \end{pmatrix}\right)_1 I_M.$$

It is easy to see that $\mathcal{P}$ has eigenvalues $\ell_1 = 1$ and $\ell_2 = 1 - \mu_\nu (M+N)/(\Lambda N)$ with eigenvectors $\underline{m}_1 = (1, 1)$ and $\underline{m}_2 = (N, -M)^T/(M+N)$. Consequently,

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{M}{N+M} \underline{m}_1 + \underline{m}_2,$$

which yields

$$\left(\mathcal{P}^k \begin{pmatrix} 1 \\ 0 \end{pmatrix}\right)_1 = \frac{M}{N+M} + \frac{N}{M+N} \left(1 - \mu_\nu \frac{M+N}{\Lambda N}\right)^k.$$

This proves Theorem 7. □

7. Boltzmann–Kac Collisions

In this section we show that the above results can also be extended, at least in a particular case, to three-dimensional Boltzmann–Kac collisions.

Again we consider a system of M particles coupled to a reservoir consisting of N particles, but now with velocities $v_1, \ldots, v_M, w_1, \ldots, w_N \in \mathbb{R}^3$. The collisions between a pair of particles have to conserve energy and momentum,

$$z_i^2 + z_j^2 = (z_i^*)^2 + (z_j^*)^2$$
$$z_i + z_j = z_i^* + z_j^*,$$

where z can be either the velocity of a system particle v or of a reservoir particle w. A convenient parametrization of the post-collisional velocities in terms of the velocities before the collision is given by

$$z_i^*(\omega) = z_i - \omega \cdot (z_i - z_j)\, \omega$$
$$z_j^*(\omega) = z_j + \omega \cdot (z_i - z_j)\, \omega, \quad \text{where } \omega \in \mathbb{S}^2.$$

This is the so-called ω-representation. This representation is particularly useful, because the velocities are related to each other by a *linear* transformation, and the strategy used to proof the results for the one-dimensional Kac system carries over rather directly. The direction ω will be chosen according to the uniform probability distribution on the unit sphere $\mathbb{S}^2$.

Introduce the operators

$$(R_{ij} f)(\boldsymbol{z}) = \int_{\mathbb{S}^2} f(r_{ij}(\omega)^{-1}\boldsymbol{z})\, \mathrm{d}\omega,$$

where $\mathrm{d}\omega$ denotes the uniform probability measure on the sphere and the matrices $r_{ij}(\omega)$ are symmetric involutions acting as

$$\begin{pmatrix} z_i^* \\ z_j^* \end{pmatrix} = \begin{pmatrix} I - \omega\omega^T & \omega\omega^T \\ \omega\omega^T & I - \omega\omega^T \end{pmatrix} \begin{pmatrix} z_i \\ z_j \end{pmatrix}$$

on the velocities of the particles i and j, and as identities otherwise. They will replace the one-dimensional Kac collision operators in (3) in the otherwise unchanged generator of the time evolution. Notice that the matrices $r_{ij}(\omega)$ are orthogonal, so that the expansion formula (12) still holds with the obvious changes in the dimension of the single-particle spaces.

We prove an analog of Theorem 1 for the case of three-dimensional Boltzmann–Kac collisions and pseudo-Maxwellian molecules.

Theorem 8. *Let $N \geq M$ and $F_0(\boldsymbol{v}, \boldsymbol{w}) = f_0(\boldsymbol{v})\, \mathrm{e}^{-\pi|\boldsymbol{w}|^2}$ for some probability distribution f_0 on $\mathbb{R}^{3M}$. Then the entropy of the marginal*

$$f(\boldsymbol{v}, t) := \int_{\mathbb{R}^{3N}} \left(\mathrm{e}^{\mathcal{L}t} F_0\right)(\boldsymbol{v}, \boldsymbol{w})\, \mathrm{d}\boldsymbol{w}$$

with respect to the thermal state $\mathrm{e}^{-\pi|\boldsymbol{v}|^2}$ is bounded by

$$S(f(\cdot, t)) \leq \left[\frac{N}{N+M} + \frac{N}{N+M} \mathrm{e}^{-\frac{\mu}{3}\frac{N+M}{N}t}\right] S(f_0)\, .$$

Remark 3. The result in three dimensions is very similar to the case of one-dimensional Kac collisions, with the difference that the rate of exponential decay is $\mu/3$ instead of μ_ρ. The factor $1/3$ comes from the fact that $\int_{\mathbb{S}^2} \mathrm{d}\omega\, \omega\omega^T = I_3/3$. It would be interesting to cover the true Maxwellian molecules interaction

$$(R_{ij} f)(z) = \int_{\mathbb{S}^2} b\left(\frac{v_i - v_j}{|v_i - v_j|} \cdot \omega\right) f(r_{ij}(\omega)^{-1} z)\, \mathrm{d}\omega\, .$$

However, the dependence of the scattering rate b on the velocities doesn't seem to be treatable with the above methods.

The proof of Theorem 8 deviates from the one-dimensional case essentially in only two places: the sum rule and the discrete approximation of the integrals. We begin by proving an analogue of Theorem 7. Most of the steps for the computation of the matrix Z in (27) are the same. What remains is to compute

$$Z := \sum_{\alpha_1,\dots,\alpha_k} \lambda_{\alpha_1} \cdots \lambda_{\alpha_k} \int_{\mathbb{S}^2\times\cdots\times\mathbb{S}^2} \mathrm{d}\omega_1 \cdots \mathrm{d}\omega_k\, A_k(\underline{\alpha}, \underline{\omega}) A_k(\underline{\alpha}, \underline{\omega})^T,$$

which is somewhat different for the case of Boltzmann–Kac collisions. Recall that $A_k(\underline{\alpha}, \underline{\omega})$ is the upper left $3M \times 3M$ block of $[\prod_{j=1}^k r_{\alpha_j}(\omega_j)]^{-1}$, i.e.,

$$A_k(\underline{\alpha}, \underline{\omega}) = P_{3M}[\Pi_{j=1}^k r_{\alpha_j}(\omega_j)]^{-1} P_{3M}^T$$

with the projection $P_{3M} = \begin{pmatrix} I_{3M} & 0 \end{pmatrix}$ from $\mathbb{R}^{3M+3N} \to \mathbb{R}^{3M}$. In particular, by linearity,

$$Z = P_{3M}\left(\sum_{\alpha_1,\ldots,\alpha_k} \lambda_{\alpha_1}\cdots\lambda_{\alpha_k}\int_{(\mathbb{S}^2)^k} \mathrm{d}\underline{\omega}\left[\prod_{j=1}^k r_{\alpha_j}(\omega_j)\right]^{-1}\begin{pmatrix} I_{3M} & 0 \\ 0 & 0\end{pmatrix}\left[\prod_{j=1}^k r_{\alpha_j}(\omega_j)\right]\right)P_{3M}^T\,.$$

As in the proof of Theorem 7 we have

Lemma 3. *Let $\alpha, \beta \geq 0$. Then*

$$\sum_{1\leq i<j\leq M+N} \lambda_{ij}\int_{\mathbb{S}^2} \mathrm{d}\omega\, r_{ij}(\omega)^{-1}\begin{pmatrix} \alpha I_{3M} & 0 \\ 0 & \beta I_{3N}\end{pmatrix} r_{ij}(\omega) = \begin{pmatrix} \alpha' I_{3M} & 0 \\ 0 & \beta' I_{3N}\end{pmatrix},$$

where α', β' are related to α, β by

$$\begin{pmatrix} \alpha' \\ \beta' \end{pmatrix} = \mathcal{P}\begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \quad \mathcal{P} = I_2 - \frac{\mu}{3\Lambda}\begin{pmatrix} 1 & -1 \\ -\frac{M}{N} & \frac{M}{N}\end{pmatrix}\,.$$

Notice that the matrix $\mathcal{P}$ of Lemma 3 has eigenvalues 1 and $1 - \mu/(3\Lambda)\,(1 + M/N)$ with corresponding eigenvectors $\begin{pmatrix} 1 & 1 \end{pmatrix}^T$ and $\begin{pmatrix} -N/M & 1 \end{pmatrix}^T$. Repeated application of Lemma 3 then implies, see also the argument in the one-dimensional case,

$$\sum_{\alpha_1,\ldots,\alpha_k} \lambda_{\alpha_1}\cdots\lambda_{\alpha_k}\int_{(\mathbb{S}^2)^k} \mathrm{d}\underline{\omega}\left[\prod_{j=1}^k r_{\alpha_j}(\omega_j)\right]^{-1}\begin{pmatrix} \alpha I_{3M} & 0 \\ 0 & \beta I_{3N}\end{pmatrix}\left[\prod_{j=1}^k r_{\alpha_j}(\omega_j)\right]$$
$$= \begin{pmatrix} \alpha^{(k)} I_{3M} & 0 \\ 0 & \beta^{(k)} I_{3N}\end{pmatrix},$$

where

$$\begin{pmatrix} \alpha^{(k)} \\ \beta^{(k)} \end{pmatrix} = \mathcal{P}^k\begin{pmatrix} \alpha \\ \beta \end{pmatrix}\,.$$

Before we prove Lemma 3, let us make an easy observation.

Corollary 3. *In the particular case $\alpha = 1$, $\beta = 0$, we get*

$$Z = \left[\frac{M}{M+N} + \frac{N}{M+N}\left(1 - \frac{\mu}{3\Lambda}\left(1 + \frac{M}{N}\right)\right)^k\right] I_{3M}.$$

Proof of Lemma 3. For $1 \leq i < j \leq M$ (respectively for $M + 1 \leq i < j \leq M + N$) the operators $r_{ij}(\omega)$ only act non-trivially in the first $3M$ (last $3N$) variables. Taking into account that $r_{ij}(\omega)^{-1}\, I\, r_{ij}(\omega) = I$, we obtain

$$\frac{\lambda_S}{M-1}\sum_{1\leq i<j\leq M}\int_{\mathbb{S}^2} \mathrm{d}\omega\, r_{ij}(\omega)^{-1}\begin{pmatrix} \alpha I_{3M} & 0 \\ 0 & \beta I_{3N}\end{pmatrix} r_{ij}(\omega) = \frac{M\lambda_S}{2}\begin{pmatrix} \alpha I_{3M} & 0 \\ 0 & \beta I_{3N}\end{pmatrix},$$

and

$$\frac{\lambda_R}{N-1} \sum_{M+1 \le i < j \le M+N} \int_{\mathbb{S}^2} \mathrm{d}\omega\, r_{ij}(\omega)^{-1} \begin{pmatrix} \alpha I_{3M} & 0 \\ 0 & \beta I_{3N} \end{pmatrix} r_{ij}(\omega) = \frac{N\lambda_R}{2} \begin{pmatrix} \alpha I_{3M} & 0 \\ 0 & \beta I_{3N} \end{pmatrix}.$$

It remains to look at the interaction terms $i = 1, \dots, M$ and $j = M+1, \dots, M+N$. Notice that

$$\begin{aligned} & r_{ij}(\omega)^{-1} \begin{pmatrix} \alpha I_{3M} & 0 \\ 0 & \beta I_{3N} \end{pmatrix} r_{ij}(\omega) \\ &= \begin{pmatrix} \alpha I_{3M} & 0 \\ 0 & \beta I_{3N} \end{pmatrix} + \left(\begin{array}{ccc|ccc} 0 & & & & & \\ & (\beta-\alpha)\omega\omega^T & & & 0 & \\ & & 0 & & & \\ \hline & & & 0 & & \\ & 0 & & & (\alpha-\beta)\omega\omega^T & \\ & & & & & 0 \end{array} \right), \end{aligned}$$

where the non-zero entries in the second summand on the right-hand side correspond to the i^{th}, respectively j^{th}, 3×3 block on the diagonal. Since $\int_{\mathbb{S}^2} \mathrm{d}\omega\, \omega\omega^T = 1/3\, I_3$, we obtain

$$\begin{aligned} & \frac{\mu}{N} \sum_{i=1}^{M} \sum_{j=M+1}^{M+N} \int_{\mathbb{S}^2} \mathrm{d}\omega\, r_{ij}(\omega)^{-1} \begin{pmatrix} \alpha I_{3M} & 0 \\ 0 & \beta I_{3N} \end{pmatrix} r_{ij}(\omega) \\ &= \mu M \begin{pmatrix} \alpha I_{3M} & 0 \\ 0 & \beta I_{3N} \end{pmatrix} + \frac{\mu}{3}(\alpha-\beta) \begin{pmatrix} -I_{3M} & 0 \\ 0 & \frac{M}{N} I_{3N} \end{pmatrix}. \end{aligned}$$

Recall the definition of $\Lambda = M\lambda_S/2 + N\lambda_R/2 + \mu M$. Hence, summation of all the three contributions yields the statement of the Lemma. □

As in the one-dimensional case, in order to apply the geometric Brascamp–Lieb inequality Theorem 6, we need to approximate the uniform probability measure $\mathrm{d}\omega$ on the sphere by a suitable sequence of discrete measures as in the one-dimensional case (see Lemma 2). Additionally, in each step of the discretization, the constraint $\int_{\mathbb{S}^2} \mathrm{d}\omega\, \omega\omega^T = 1/3 I$, has to hold. This is important, because it guarantees that the geometric Brascamp–Lieb condition, i.e., the sum rule (21), holds in each step.

In order to find such an approximation, we parametrize the sphere in the usual way by spherical coordinates

$$\omega = \omega(\theta, \varphi) = \begin{pmatrix} \sin\theta\cos\varphi \\ \sin\theta\sin\varphi \\ \cos\theta \end{pmatrix}$$

for $\theta \in [0, \pi]$ and $\varphi \in [0, 2\pi]$. For $K, L \in \mathbb{N}$ we introduce the measures

$$\Phi_K := \frac{\pi}{K} \sum_{j=0}^{2K-1} \delta_{\frac{\pi}{K} j} \quad \text{on } [0, 2\pi], \quad \text{and}$$

$$\Theta_L := \sum_{i=1}^{L} \frac{2}{(1-u_i^2)^{3/2} (P_L'(u_i))^2} \delta_{\arccos u_i} \quad \text{on } [0, \pi],$$

where P_L is the Legendre polynomial of order L on $[-1, 1]$, and u_i, $i = 1, \ldots, L$, are its zeros. Then, if $f \in \mathcal{C}[0, 2\pi]$ and $g \in \mathcal{C}[-1, 1]$,

$$\int_0^{2\pi} f(\varphi)\, \Phi_K(\mathrm{d}\varphi) = \frac{\pi}{K} \sum_{j=0}^{2K-1} f\left(\frac{\pi}{K} j\right) \to \int_0^{2\pi} f(\varphi)\, \mathrm{d}\varphi$$

as $K \to \infty$ as Riemann sum. Furthermore,

$$\int_0^{\pi} g(\cos\theta)\, \sin\theta\, \Theta_L(\mathrm{d}\theta) = \sum_{i=1}^{L} \frac{2\sin(\arccos u_i)}{(1-u_i^2)^{3/2}(P_L'(u_i))^2} g(u_i)$$
$$= \sum_{i=1}^{L} \frac{2}{(1-u_i^2)(P_L'(u_i))^2} g(u_i) \longrightarrow \int_{-1}^{1} g(u)\, \mathrm{d}u = \int_0^{\pi} g(\cos\theta)\, \sin\theta\, \mathrm{d}\theta$$

as $L \to \infty$ by Gauss-Legendre quadrature . The latter approximation is exact for polynomials of order less or equal to $2L-1$. In particular, we have

$$\int_0^{\pi} \cos^2\theta\, \sin\theta\, \Theta_L(\mathrm{d}\theta) = \int_0^{\pi} \cos^2\theta\, \sin\theta\, \mathrm{d}\theta = \frac{2}{3}, \text{ and}$$
$$\int_0^{\pi} \sin^3\theta\, \Theta_L(\mathrm{d}\theta) = \int_0^{\pi} \sin^3\theta\, \mathrm{d}\theta = \frac{4}{3},$$

for all $L \geq 2$. It is easy to check that

$$\int_0^{2\pi} \sin\varphi \cos\varphi\, \Phi_K(\mathrm{d}\varphi) = 0,$$
$$\int_0^{2\pi} \sin\varphi\, \Phi_K(\mathrm{d}\varphi) = \int_0^{2\pi} \cos\varphi\, \Phi_K(\mathrm{d}\varphi) = 0,$$
$$\int_0^{2\pi} \sin^2\varphi\, \Phi_K(\mathrm{d}\varphi) = \int_0^{2\pi} \cos^2\varphi\, \Phi_K(\mathrm{d}\varphi) = \pi,$$

for all $K \geq 2$. Consequently,

$$\frac{1}{4\pi} \int_0^{2\pi} \omega(\theta,\varphi)\omega(\theta,\varphi)^T\, \Theta_L(\mathrm{d}\theta)\Phi_K(\mathrm{d}\varphi) = \frac{1}{3} I_3$$

for all $K, L \geq 2$. It follows that Z is not changed by replacing the uniform measure on $\mathbb{S}^2$ by the above discrete approximation, in particular, Z is still proportional to the identity matrix, which guarantees the applicability of the geometric Brascamp–Lieb inequality.

This concludes the proof of Theorem 8. □

Acknowledgements. A.G. and T.R. would like to thank Georgia Tech for its hospitality. The work of M.L. and A.G. was supported in part by NSF grant DMS-1600560 and the Humboldt Foundation. A.G. and T.R. gratefully acknowledge financial support by the Deutsche Forschungsgemeinschaft (DFG) through GRK 1838 (A.G.) and CRC 1173 (T.R.). F.B. gratefully acknowledges financial support from the Simons Foundation award number 359963. T.R. thanks the Karlsruhe House of Young Scientists (KHYS) for a Research Travel Grant supporting the stay at Georgia Tech.

A The Geometric Brascamp–Lieb Inequality and the Entropy Inequality

In this section we prove Theorem 6. We use the same strategy as in [13] and [5] which consists of transporting the functions f_i with the heat kernel in such a way that the right-hand side of (22) remains fixed while the left-hand side of that inequality increases. The results in [5] are quite general but for the special case in which the sum rule (21) holds, the proof is quite simple and this is one of the reasons why we include it here.

Proof of Theorem 6. The inequality (22) is equivalent to

$$\int_{\mathbb{R}^M} \prod_{i=1}^{K} f_i^{c_i}(B_i \mathbf{v})\, \mathrm{d}\mathbf{v} \le \prod_{i=1}^{K} \left(\int_{H_i} f_i(u)\, \mathrm{d}u \right)^{c_i}. \tag{30}$$

This follows from the identity

$$\prod_{i=1}^{K} \left(e^{-\pi |B_i \mathbf{v}|^2} \right)^{c_i} = e^{-\pi \sum_{i=1}^{K} (\mathbf{v},\, c_i B_i^T B_i \mathbf{v})} = e^{-\pi |\mathbf{v}|^2}.$$

We transport the functions f_i by the heat flow, that is we define

$$f_i(B_i \mathbf{v}, t) := \frac{1}{(4\pi t)^{M/2}} \int_{\mathbb{R}^M} e^{-|\mathbf{v}-\mathbf{w}|^2/(4t)} f_i(B_i \mathbf{w}) \mathrm{d}\mathbf{w}. \tag{31}$$

For the above definition to make sense, we have to show that the right-hand side is a function of $B_i \mathbf{v}$ alone. The condition $B_i B_i^T = I_{H_i}$ means that the matrix $P_i = B_i^T B_i$ is an orthogonal projection onto a d_i dimensional subspace of $\mathbb{R}^M$. Moreover, $B_i P_i = I_{H_i} B_i = B_i$. We rewrite the integral (31) by splitting it in an integral over $\mathbf{w}' \in \mathrm{Ran}\ P_i$ and one over integration over $\mathbf{w}'' \in \mathrm{Ran}\ P_i^{\perp}$. Carrying out the integration over $\mathbf{w}''$ we obtain

$$\begin{aligned} f_i(B_i \mathbf{v}, t) &= \frac{1}{(4\pi t)^{M/2}} \int_{\mathrm{Ran}\ P_i} \int_{\mathrm{Ran}\ P_i^{\perp}} e^{-|(P_i \mathbf{v} - P_i \mathbf{w}')|^2/(4t)} \\ &\quad e^{-|(P_i^{\perp} \mathbf{v} - \mathbf{w}'')|^2/(4t)} f_i(B_i P_i \mathbf{w})\, \mathrm{d}\mathbf{w}' \mathrm{d}\mathbf{w}'' \\ &= \frac{1}{(4\pi t)^{d_i/2}} \int_{\mathrm{Ran}\ P_i} e^{-|(P_i \mathbf{v} - P_i \mathbf{w}')|^2/(4t)} f_i(B_i P_i \mathbf{w}')\, \mathrm{d}\mathbf{w}' \\ &= \frac{1}{(4\pi t)^{d_i/2}} \int_{\mathrm{Ran}\ P_i} e^{-|(B_i \mathbf{v} - B_i \mathbf{w}')|^2/(4t)} f_i(B_i \mathbf{w}')\, \mathrm{d}\mathbf{w}' \\ &= \frac{1}{(4\pi t)^{d_i/2}} \int_{H_i} e^{-|(B_i \mathbf{v} - u)|^2/(4t)} f_i(u)\, \mathrm{d}u \end{aligned}$$

where, in the last equality, we have used that B_i maps the range of P_i isometrically onto H_i. This justifies (31). Moreover, the above computation also shows that

$$\int_{H_i} f_i(u, t) \mathrm{d}u = \int_{H_i} f_i(u) \mathrm{d}u$$

so that the right-hand side of the inequality (30) does not change under the heat flow.

We now show that the left-hand side of (30) is an increasing function of t. It is convenient to set $\phi_i(u,t) = \log f_i(u,t)$. Differentiating the function $\phi_i(B_i\mathbf{v},t)$ with respect to t yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\phi_i(B_i\mathbf{v},t) = \Delta_{\mathbf{v}}\phi_i(B_i\mathbf{v},t) + |\nabla_{\mathbf{v}}\phi_i(B_i\mathbf{v},t)|^2.$$

Moreover,

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^M}\prod_{i=1}^K f_i^{c_i}(B_i\mathbf{v},t)\,\mathrm{d}\mathbf{v}$$
$$= \sum_{m=1}^K c_m\int_{\mathbb{R}^M}[\Delta_{\mathbf{v}}\phi_m(B_m\mathbf{v},t) + |\nabla_{\mathbf{v}}\phi_m(B_m\mathbf{v},t)|^2]\prod_{i=1}^K f_i^{c_i}(B_i\mathbf{v},t)\,\mathrm{d}\mathbf{v}.$$

Integrating by parts the term containing the Laplacian yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^M}\prod_{i=1}^K f_i^{c_i}(B_i\mathbf{v},t)\,\mathrm{d}\mathbf{v}$$
$$= \sum_{m=1}^K c_m\int_{\mathbb{R}^M}|\nabla_{\mathbf{v}}\phi_m(B_m\mathbf{v},t)|^2\prod_{i=1}^K f_i^{c_i}(B_i\mathbf{v},t)\mathrm{d}\mathbf{v}$$
$$- \sum_{m,\ell=1}^K c_m c_\ell\int_{\mathbb{R}^M}\nabla_{\mathbf{v}}\phi_m(B_m\mathbf{v},t)\cdot\nabla_{\mathbf{v}}\phi_\ell(B_\ell\mathbf{v},t)\prod_{i=1}^K f_i^{c_i}(B_i\mathbf{v},t)\mathrm{d}\mathbf{v}.$$

Finally, using that

$$\nabla_{\mathbf{v}}\phi_m(B_m\mathbf{v},t) = B_m^T(\nabla\phi_m)(B_m\mathbf{v})$$

we get

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^M}\prod_{i=1}^K f_i^{c_i}(B_i\mathbf{v},t)\,\mathrm{d}\mathbf{v}$$
$$= \sum_{m=1}^K c_m\int_{\mathbb{R}^M}|B_m^T(\nabla\phi_m)(B_m\mathbf{v},t)|^2\prod_{i=1}^K f_i^{c_i}(B_i\mathbf{v},t)\mathrm{d}\mathbf{v}$$
$$- \sum_{m,\ell=1}^K c_m c_\ell\int_{\mathbb{R}^M}B_m^T(\nabla\phi_m)(B_m\mathbf{v},t)\cdot B_\ell^T(\nabla\phi_\ell)(B_\ell\mathbf{v},t)\prod_{i=1}^K f_i^{c_i}(B_i\mathbf{v},t)\mathrm{d}\mathbf{v}.$$

We claim that this expression is non-negative. The vectors $\nabla\phi_m \in H_m$ are arbitrary and hence the problem is reduced to proving that for any set of vectors $V_m \in H_m$, $m = 1,\ldots,K$, it holds

$$\sum_{m=1}^K c_m|B_m^T V_m|^2 - \sum_{m,\ell=1}^K c_m c_\ell B_m^T V_m\cdot B_\ell^T V_\ell \ \geq\ 0.$$

Recalling that $B_m B_m^T = I_{H_m}$ and setting $Y = \sum_\ell c_\ell B_\ell^T V_\ell$ we conclude that it is enough to show that

$$|Y|^2 \le \sum_{m=1}^K c_m |V_m|^2.$$

This follows easily, since, by applying Schwarz's inequality, we find that

$$|Y|^2 = \sum_{\ell=1}^K c_\ell Y \cdot B_\ell^T V_\ell = \sum_{\ell=1}^K c_\ell B_\ell Y \cdot V_\ell \le \left(\sum_{\ell=1}^K c_\ell |B_\ell Y|^2\right)^{1/2} \left(\sum_{\ell=1}^K c_\ell |V_\ell|^2\right)^{1/2}.$$

Combining this with (21), we learn that

$$|Y|^2 \le \left(Y \cdot \sum_{\ell=1}^K c_\ell B_\ell^T B_\ell Y\right)^{1/2} \left(\sum_{\ell=1}^K c_\ell |V_\ell|^2\right)^{1/2} = |Y| \left(\sum_{\ell=1}^K c_\ell |V_\ell|^2\right)^{1/2}.$$

Thus we have that, when applying (30) to the functions $f_i(u,t)$, the left hand side is an increasing function of t while the right hand side does not depends on t. It is thus enough to show that the inequality holds for large t. Using once more the sum-rule (21), we see that

$$\begin{aligned}
&\int_{\mathbb{R}^M} \prod_{i=1}^K \frac{1}{(4\pi t)^{c_i d_i/2}} \left[\int_{H_i} e^{-\frac{|B_i \mathbf{v} - u|^2}{4t}} f_i(u) du\right]^{c_i} d\mathbf{v} \\
&= \frac{1}{(4\pi)^{M/2}} \int_{\mathbb{R}^M} \prod_{i=1}^K \left[\int_{H_i} e^{-\frac{|B_i \mathbf{v} - t^{-1/2} u|^2}{4}} f_i(u) du\right]^{c_i} d\mathbf{v} \\
&\xrightarrow{t\to\infty} \frac{1}{(4\pi)^{M/2}} \int_{\mathbb{R}^M} e^{-\frac{|\mathbf{v}|^2}{4}} \prod_{i=1}^K \left[\int_{H_i} f_i(u) du\right]^{c_i} d\mathbf{v} = \prod_{i=1}^K \left[\int_{H_i} f_i(u) du\right]^{c_i}
\end{aligned}$$

which proves the first part of Theorem 6.

To prove the entropy inequality (23) we follow [11]. Let h be a non-negative function whose L^1 norm is one and whose entropy is finite. An elementary computation then shows that

$$\begin{aligned}
&\int_{\mathbb{R}^M} h(\mathbf{v}) \log h(\mathbf{v})\, e^{-\pi|\mathbf{v}|^2} \mathrm{d}\mathbf{v} \\
&\quad = \sup_{\Phi} \left\{ \int_{\mathbb{R}^M} h(\mathbf{v}) \Phi(\mathbf{v})\, e^{-\pi|\mathbf{v}|^2}\, \mathrm{d}\mathbf{v} - \log \int_{\mathbb{R}^M} e^{\Phi(\mathbf{v})} e^{-\pi|\mathbf{v}|^2}\, \mathrm{d}\mathbf{v} \right\}.
\end{aligned}$$

Now, we set

$$\Phi(\mathbf{v}) = \sum_{i=1}^K c_i \log f_i(B_i \mathbf{v}).$$

This leads to the lower bound

$$\begin{aligned}
&\int_{\mathbb{R}^M} h(\mathbf{v}) \log h(\mathbf{v})\, e^{-\pi|\mathbf{v}|^2}\, \mathrm{d}\mathbf{v} \\
&\geq \sum_{i=1}^{K} c_i \int_{\mathbb{R}^M} h(\mathbf{v}) \log f_i(B_i\mathbf{v})\, e^{-\pi|\mathbf{v}|^2}\, \mathrm{d}\mathbf{v} - \log \int_{\mathbb{R}^M} \prod_{i=1}^{K} f_i(B_i\mathbf{v})^{c_i}\, e^{-\pi|\mathbf{v}|^2}\, \mathrm{d}\mathbf{v} \\
&\geq \sum_{i=1}^{K} c_i \int_{\mathbb{R}^M} h(\mathbf{v}) \log f_i(B_i\mathbf{v})\, e^{-\pi|\mathbf{v}|^2}\, \mathrm{d}\mathbf{v} - \log \left[\prod_{i=1}^{K} \left(\int_{H_i} f_i(u)\, e^{-\pi|u|^2}\, \mathrm{d}u \right)^{c_i} \right],
\end{aligned}$$

where the second step is a consequence of the Brascamp–Lieb inequality (22). □

B Proof of Lemma 2

Proof. For K any positive integer we convolve ρ with the non-negative trigonometric polynomial

$$p_K(\theta) := \frac{1}{2K+1} \left(\sum_{k=-K}^{K} e^{ik\theta} \right)^2 = \sum_{m=-2K}^{2K} \left(1 - \frac{|m|}{2K+1} \right) e^{im\theta},$$

and obtain a probability density ρ_K. The Fourier coefficients of ρ_K are given by

$$\widehat{\rho}_K(m) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \rho_K(\theta)\, e^{-im\theta}\, \mathrm{d}\theta = \widehat{\rho}(m) \left(1 - \frac{|m|}{2K+1} \right)$$

for $|m| \leq 2K$ and are zero otherwise. In particular,

$$\begin{aligned}
\frac{1}{2\pi} \int_{-\pi}^{\pi} \rho_K(\theta) \sin\theta \cos\theta\, \mathrm{d}\theta &= \frac{\widehat{\rho}_K(-2) - \widehat{\rho}_K(2)}{4i} = \left(1 - \frac{2}{2K+1} \right) \frac{\widehat{\rho}(-2) - \widehat{\rho}(2)}{4i} \\
&= \left(1 - \frac{2}{2K+1} \right) \frac{1}{2\pi} \int_{-\pi}^{\pi} \rho(\theta) \sin\theta \cos\theta\, \mathrm{d}\theta = 0.
\end{aligned}$$

With ρ_K we construct the measure

$$\nu_K(d\theta) = \frac{2\pi}{4K+1} \sum_{\ell=-2K}^{2K} \rho_K \left(\frac{2\pi\ell}{4K+1} \right) \delta \left(\theta - \frac{2\pi\ell}{4K+1} \right) \mathrm{d}\theta.$$

The measure ν_K is positive since $\rho_K((2\pi\ell)/(4K+1)) \geq 0$. Moreover, for all $m \in \mathbb{Z}$ with $|m| \leq 2K$ the Fourier coefficients $\widehat{\nu}_K(m)$ and $\widehat{\rho}_K(m)$ coincide. In particular, we have

$$\int_{-\pi}^{\pi} \nu_K(\mathrm{d}\theta) \sin\theta \cos\theta = 0\,.$$

To see this, we compute

$$\widehat{\nu}_K(m) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \nu_K(\theta) e^{-im\theta}\, \mathrm{d}\theta = \frac{1}{4K+1} \sum_{\ell=-2K}^{2K} \rho_K \left(\frac{2\pi\ell}{4K+1} \right) e^{-2\pi im\ell/(4K+1)}$$

for $|m| \leq 2K$. Observe that

$$\rho_K\left(\frac{2\pi\ell}{4K+1}\right) = \sum_{k=-2K}^{2K} \widehat{\rho}_K(k) e^{2\pi i k\ell/(4K+1)},$$

and, as a consequence,

$$\widehat{\nu}_K(m) = \frac{1}{4K+1} \sum_{\ell=-2K}^{2K} \sum_{k=-2K}^{2K} \widehat{\rho}_K(k) e^{2\pi i \ell(k-m)/(4K+1)}.$$

But

$$\sum_{\ell=-2K}^{2K} e^{2\pi i \ell(k-m)/(4K+1)} = \begin{cases} 4K+1 & \text{if } k = m \\ 0 & \text{if } k \neq m, \end{cases}$$

and hence we conclude that

$$\widehat{\nu}_K(m) = \widehat{\rho}_K(m) \tag{32}$$

for $|m| \leq 2K$. It is easy to see that for any continuous function f on $[-\pi, \pi]$,

$$\lim_{K\to\infty} \int_{-\pi}^{\pi} f(\theta)\nu_K(\mathrm{d}\theta) = \int_{-\pi}^{\pi} f(\theta)\rho(\theta)\,\mathrm{d}\theta.$$

□

References

1. Ball, K.: Volumes of sections of cubes and related problems. In: Geometric Aspects of Functional Analysis: Israel Seminar (GAFA) 1987–88, Volume **1376** of the Series Lecture Notes in Mathematics, pp 251–260. Springer, Berlin (1989)
2. Ball, K.: Volume ratios and a reverse isoperimetric inequality,. J. Lond. Math. Soc. (Second Series) **44**, 351–359 (1991).
3. Barthe, F.: On a reverse form of the Brascamp–Lieb inequality. Inventiones Mathematicae **134**, 335–361 (1998)
4. Barthe, F.: A continuous version of the Brascamp–Lieb inequalities. In: Geometric Aspects of Functional Analysis: Israel Seminar 2002–2003, Volume **1850** of the Series Lecture Notes in Mathematics, pp 53–63. Springer, Berlin (2004)
5. Bennett, J., Carbery, A., Christ, M., Tao, T.: The Brascamp–Lieb inequalities: finiteness, structure and extremals. Geom. Funct. Anal. **17**, 1343–1415 (2008)
6. Bonetto, F., Loss, M., Tossounian, H., Vaidyanathan, R.: Uniform approximation of a Maxwellian thermostat by finite reservoirs. Commun. Math. Phys. **351**, 311–339 (2017)
7. Bonetto, F., Loss, M., Vaidyanathan, R.: The Kac model coupled to a thermostat. J. Stat. Phys. **156**, 647–667 (2014)
8. Brascamp, H.J., Lieb, E.H.: Best constants in Young's inequality, its converse, and its generalization to more than three functions. Adv. Math. **20**, 151–173 (1976)
9. Carlen, E.A., Carvalho, M.C., Loss, M.: Many-body aspects of approach to equilibrium. In: Journées "Équations aux Dérivées Partielles" (La Chapelle sur Erdre, 2000), Exp. No. XI, 12 pp., Université de Nantes, Nantes, (2000)
10. Carlen, E.A., Carvalho, M.C., Loss, M.: Determination of the spectral gap for Kac's master equation and related stochastic evolution. Acta Mathematica **191**, 1–54 (2003)
11. Carlen, E.A., Cordero-Erausquin, D.: Subadditivity of the entropy and its relation to Brascamp–Lieb type inequalities. Geom. Funct. Anal. **19**, 373–405 (2009)
12. Carlen, E.A., Loss, M.: Extremals of functionals with competing symmetries. J. Funct. Anal. **88**, 437–456 (1990)

13. Carlen, E.A., Lieb, E.H., Loss, M.: A sharp analog of Young's inequality on S^N and related entropy inequalities. J. Geom. Anal. **14**, 487–520 (2004)
14. Einav, A.: On Villani's conjecture concerning entropy production for the Kac master equation. Kinet. Relat. Models **4**, 479–497 (2011)
15. Federbush, P.: Partially alternate derivation of a result of Nelson. J. Math. Phys. **10**, 50–52 (1969)
16. Gabetta, E., Toscani, G., Wennberg, B.: Metrics for probability distributions and the trend to equilibrium for solutions of the Boltzmann equation. J. Stat. Phys. **81**, 901–934 (1995)
17. Gross, L.: Logarithmic Sobolev inequalities. Am. J. Math. **97**, 1061–1083 (1975)
18. Gross, L.: Logarithmic Sobolev inequalities and contractivity properties of semigroups. In Dirichlet Forms (Varenna, 1992), Volume **1563** of the series Lecture Notes in Mathematics, pp 54–88. Springer Berlin Heidelberg, (1993)
19. Han, T.S.: Nonnegative entropy measures of multivariate symmetric correlations. Inf. Control. **36**, 133–156 (1978)
20. Janvresse, E.: Spectral gap for Kac's model of Boltzmann equation. Ann. Probab. **29**, 288–304 (2001)
21. Kac, M.: Foundations of kinetic theory. In: Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, vol. III, pp. 171–197. University of California Press, Berkeley (1956)
22. Lieb, E.H.: Gaussian kernels have only Gaussian maximizers. Inventiones Mathematicae **102**, 179–208 (1990)
23. Loomis, L.H., Whitney, H.: An inequality related to the isoperimetric inequality. Bull. Am. Math. Soc. **55**, 961–962 (1949)
24. Maslen, D.K.: The eigenvalues of Kac's master equation. Mathematische Zeitschrift **243**, 291–331 (2003)
25. Mischler, S., Mouhot, C.: About Kac's program in kinetic theory. Comptes Rendus Mathématique (Académie Des Sciences Paris). **349**, 1245–1250 (2011)
26. Mischler, S., Mouhot, C.: Kac's program in kinetic theory. Inventiones Mathematicae **193**, 1–147 (2013)
27. Nelson, E.: The free Markoff field. J. Funct. Anal. **12**, 211–227 (1973)
28. Tossounian, H., Vaidyanathan, R.: Partially thermostated Kac model. J. Math. Phys. **56**, 083301 (2015)
29. Villani, C.: Cercignani's conjecture is sometimes true and always almost true. Commun. Math. Phys. **234**, 455–490 (2003)

Communicated by C. Mouhot