Journal of Statistical Physics (2022) 188:7
https://doi.org/10.1007/s10955-022-02932-4

Grand Canonical Evolution for the Kac Model

Justin Beck[1] · Federico Bonetto[2]

Received: 10 April 2021 / Accepted: 1 May 2022 / Published online: 21 May 2022

Abstract
We study a model of random colliding particles interacting with an infinite reservoir at fixed temperature and chemical potential. Interaction between the particles is modeled via a Kac master equation (in: Kay, Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, University of California Press, Berkeley and Los Angeles, 1956). Moreover, particles can leave the system toward the reservoir or enter the system from the reservoir. The system admits a unique steady state given by the Grand Canonical Ensemble at temperature $T = \beta^{-1}$ and chemical potential χ. We show that any initial state converges exponentially fast to equilibrium by computing the spectral gap of the generator in a suitable L^2 space and by showing exponential decrease of the relative entropy with respect to the steady state. We also show propagation of chaos and thus the validity of a Boltzmann-Kac type equation for the particle density in the infinite system limit.

Keywords Kac model · Approach to equilibrium · Particle reservoir

1 Introduction

In 1955, Mark Kac [14] introduced a simple model to study the evolution of a dilute gas of N particles with unit mass undergoing pairwise collisions. Instead of following the deterministic evolution of the particles until a collision takes place, he considered particles that collide at random times with every particle undergoing, on average, a given number of collisions per unit time. Moreover, when a collision takes place, the energy of the two particles is randomly redistributed between them. In such a situation, one can neglect the position of the particles

Communicated by Eric A. Carlen.

✉ Federico Bonetto
bonetto@math.gatech.edu

Justin Beck
jbeck6@nd.edu

1 Department of Mathematics, University of Notre Dame, 255 Hurley Bldng, Notre Dame, IN 46556, USA

2 School of Mathematics, Georgia Institute of Technology, 686 Cherry St., Atlanta, GA 30332, USA

and focus on their velocities. To obtain a model as simple as possible, he considered particles that move in one spatial dimension. This leads to an evolution governed by a master equation for the probability distribution $f(\underline{v}_N)$, where $\underline{v}_N \in \mathbb{R}^N$ are the velocities of the particles. Since collisions preserve the kinetic energy of the system, to obtain ergodicity one has to restrict the evolution to $\underline{v}_N \in \mathbb{S}^{N-1}(\sqrt{2eN})$, that is on the surface of constant kinetic energy with e the kinetic energy per particle. To further simplify the model, he neglected the dependence of a particle collision rate on its speed, a situation sometime referred as *Maxwellian particles*. In this setting, the dynamical properties of the evolution do not depend on e and it is thus natural to set $e = 1/2$, see [14–16] for more details.

The study of the Kac master equation has been very useful to clarify and investigate notions and conjectures arising from the kinetic theory of diluted gases. We refer the reader to Kac's original works [14] and [15] for extensive discussion.

Kac's master equation also provides a natural setting to study approach to equilibrium. In the case of the standard Kac model [14], equilibrium is represented by the uniform distribution on the surface of given kinetic energy. Uniform convergence in the sense of the L^2 gap was conjectured by Kac and it was established in [13] while the gap was explicitly computed in [5].

A more natural way to define approach to equilibrium is via the relative entropy. This provides a better setting since the relative entropy, in general, grows only linearly with the number of particles. There is no result of exponential decay of relative entropy with a rate that is uniform in N for the original Kac model. Moreover, estimates of the entropy production rate seem to point to a slow decay, at least for short times, see [7, 19].

In [4], the authors studied the evolution of a dilute gas of N particles brought to equilibrium via a Maxwellian thermostat, i.e. an infinite heat reservoir at fixed temperature $T = \beta^{-1}$. The velocities of the particles in the system evolve according to the standard Kac collision process described above. On top of this, particles in the system collide with particles in the thermostat at randomly distributed times. In this way, the system and the reservoir exchange energy, but there is no exchange of particles. In particular, the kinetic energy of the system is no more preserved. They proved that the system admits as a unique steady state the Canonical Ensemble, i.e. in the steady state the probability distribution $f(\underline{v}_N)$ is the Maxwellian distribution at temperature T. Moreover, the steady state is approached exponentially fast and uniformly in N, both in the sense of the spectral gap, in a suitable L^2 space, and in the sense of the relative entropy. In both cases, the rate of approach is determined by the interaction with the thermostat while the rate of collision between particles in the system appears only in the second spectral gap. They also adapted McKean's proof [16] of propagation of chaos and obtained a Boltzmann-Kac type effective equation for the evolution of the one particle marginal in the limit $N \to \infty$.

In the present work, we study a different way to bring the system to equilibrium. As in [4], we study a system of N particles evolving through pair collisions and interacting with an infinite reservoir at given temperature T; however, the system and the reservoir are allowed to exchange particles. The evolution of the the velocities of the particles in the system is again described by a standard Kac collision process. On top of these, at random times a particle in the system can leave it while, still at random times, a particle can enter the system from the reservoir with its velocity distributed according to the Maxwellian at temperature T. Since the reservoir is infinite, no particle can enter or leave the system more than once. Clearly, in this new setting, energy and number of particles are not preserved. We show that this new evolution admits as its unique steady state the Grand Canonical Ensemble. This means that, in the steady state, the probability that the system contains N particles is given by a Poisson distribution while the probability distribution on the velocities, given the number of particles, is the Maxwellian at temperature T.

We also study the approach to equilibrium in a suitable L^2 space and in relative entropy. In both cases, we show that the rate of approach is uniform in the average number of particles. As in [4], the approach to equilibrium, both in L^2 and in relative entropy, is driven by the thermostat alone while the second spectral gap depends on the rate of binary particle collisions. Finally, we look at the emergence of an effective evolution for the particle density in the limit of a large system, that is when the average number of particles goes to infinity. This requires some adaptation of the concept of propagation of chaos since the number of particles in the system is not constant. Adapting the proof in [16], we show that the relative particle density, defined in (19) and (22) below, satisfies a Boltzmann-Kac type of equation.

The rest of the paper is organized as follows. In Sect. 2, we present the model and state our main results. Section 3 contains the proofs of our main results, while in Sect. 4 we report some open problems and present possible areas of future work. Finally the appendix contains the proofs of some technical Lemmas used in Sect. 3.

2 Model and Results

Since we want to describe a dilute gas with uniform density exchanging particles with an infinite reservoir, it is natural to assume that, in a given time, each particle in the system has the same probability of leaving it independently from the total number N of particles in the system. This implies that the flow of particles from the system to the reservoir is proportional to N. On the other hand, the probability of a particle to enter the system from the reservoir depends only on the characteristics of the reservoir, and not on N, so that the flow of particles in the system is independent from N. Finally, since the gas is dilute, given two particles in the system, their probability of colliding in a given time does not depend on the total number of particles in the system. Thus we expect the number of binary collisions in the system, in a given time, to be proportional to $\binom{N}{2}$. These are the main heuristic considerations that lead to the formulation of our model to be introduced formally below.

We consider a system of particles in one space dimension interacting with an infinite reservoir with which it exchanges particles. Since the number of particles in the system is not constant, the phase space is given by $\mathcal{R} = \bigcup_{N=0}^{\infty} \mathbb{R}^N$, where $\mathbb{R}^0 = \{\emptyset\}$ represents the state where no particle is in the system.

The evolution of the system is governed by three separate random processes. First, at exponentially distributed times a particle is added to the system with a velocity randomly chosen from a Maxwellian distribution at temperature T. To simplify notation we chose $T^{-1} = 2\pi$. Second, also at exponentially distributed times, a particle is chosen at random to exit the system and disappear forever with no chance of reentry. Finally, a pair of particles in the system is selected at random to undergo a standard Kac collision.

More precisely, let $L^1_s(\mathcal{R}) = \bigoplus_{N=0}^{\infty} L^1_s(\mathbb{R}^N)$ be the Banach space of all states $\mathbf{f} = (f_N)_{N=0}^{\infty}$, with $f_N(\underline{v}_N)$ symmetric under permutation of the v_i, defined by the norm $\|\mathbf{f}\|_1 := \sum_N \|f_N\|_{1,N}$, where $\|f_N\|_{1,N} = \int d\underline{v}_N |f_N(\underline{v}_N)|$. We say that $\mathbf{f}$ is positive if $f_N(\underline{v}_N) \geq 0$ for every N and almost every $\underline{v}_N$. If $\mathbf{f}$ is positive and $\|\mathbf{f}\|_1 = 1$ then $\mathbf{f}$ is a probability distribution on $\mathcal{R}$. In this case, for $N > 0$, $f_N(\underline{v}_N)$ represents the probability of finding N particles in the system with velocities $\underline{v}_N = (v_1, \ldots, v_N)$ while $f_0 \in \mathbb{R}$ is the probability that the system contains no particles.

The master equation for the evolution is given by

$$\frac{d}{dt}\mathbf{f} = \mathcal{L}[\mathbf{f}] := \mu(\mathcal{I}[\mathbf{f}] - \mathbf{f}) + \rho(\mathcal{O}[\mathbf{f}] - \mathcal{N}[\mathbf{f}]) + \tilde{\lambda}\mathcal{K}[\mathbf{f}] \tag{1}$$

where $\mathcal{I}$ is the *in* operator that represents the effect of introducing a particle into the system and, after symmetrization, is given by

$$(\mathcal{I}\mathbf{f})_N(\underline{v}) = \frac{1}{N}\sum_{i=1}^{N} e^{-\pi v_i^2} f_{N-1}(v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_N) \tag{2}$$

while $\mathcal{O}$ is the *out* operator that represents the effect of a random particle leaving the system

$$(\mathcal{O}\mathbf{f})_N(\underline{v}) = \sum_{i=1}^{N+1} \int dw f_{N+1}(v_1, \dots, v_{i-1}, w, v_i, \dots, v_N) \tag{3}$$

and

$$(\mathcal{N}\mathbf{f})_N(\underline{v}) = N f_N(v_1, \dots, v_N)\,.$$

Observe that, due to the symmetry of f_{N+1}, we can write

$$(\mathcal{O}\mathbf{f})_N(\underline{v}_N) = (N+1)\int dv_{N+1} f_{N+1}(\underline{v}_{N+1})\,.$$

We also define the *thermostat* operator $\mathcal{T}$ as

$$\mathcal{T} := \mu(\mathcal{I} - \mathrm{Id}) + \rho(\mathcal{O} - \mathcal{N})\,. \tag{4}$$

These definitions imply that, in every time interval dt, there is a probability μdt of a particle being added to the system. This probability is independent of the number of particles already in the system. In the same time interval, every particle in the system has a probability ρdt of leaving the system, which is, again, independent of the number of particles in the system. Thus, as discussed at the beginning of this section, the outflow of particles is proportional to N while the inflow does not depend on N.

Finally $\mathcal{K}$ represents the effect of the collisions among particles. It acts independently on each of the N particles subspaces, that is it is $(\mathcal{K}\mathbf{f})_N = K_N f_N$ with

$$K_N f_N := \sum_{1\le i<j\le N} (R_{i,j} - \mathrm{Id}) f_N := Q_N f_N - \binom{N}{2} f_N \tag{5}$$

where $R_{i,j}$ represents the effect of a collision between particles i and j:

$$(R_{i,j} f_N)(\underline{v}_N) = \frac{1}{2\pi}\int f_N(\dots, v_i\cos\theta - v_j\sin\theta, \dots, v_i\sin\theta + v_j\cos\theta, \dots)d\theta\,, \tag{6}$$

that is, $R_{i,j} f_N$ is the average of f_N over all rotations in the plane (v_i, v_j). In this way, the probability that two given particles suffer a collision in an interval dt is proportional to $\tilde{\lambda}$ and does not depend on the number of particles in the system.

Since $\mathcal{L}$ is a sum of unbounded operators that do not commute, we first need to show that (1) defines an evolution on $L^1_s(\mathcal{R})$ and that such an evolution preserves probability distributions. Observe that, notwithstanding $\mathcal{L}$ is unbounded, the operator $\mathcal{L}_N\mathbf{f}$, defined by $\mathcal{L}_N\mathbf{f} := (\mathcal{L}\mathbf{f})_N$, is bounded as an operator from $L^1_s(\mathcal{R})$ to $L^1_s(\mathbb{R}^N)$ with $\|\mathcal{L}_N\|_{1,N} \le 2\mu + (2N+1)\rho + \tilde{\lambda}N^2$. Thus we will take $D^1 = \{\mathbf{f} \mid \sum_N N^2\|f_N\|_{1,N} < \infty\}$ as the domain of $\mathcal{L}$. It is easy to see that D^1 is dense in $L^1_s(\mathcal{R})$.

In Sect. 3.1 we will build a semigroup of continuous operators $e^{t\mathcal{L}}$ that solves (1) for initial data $\mathbf{f} \in D^1$ and show that $e^{t\mathcal{L}}$ preserves probability distributions.

Lemma 1 *There exists a semigroup of continuous operators $e^{t\mathcal{L}}$ such that if $\mathbf{f} \in D^1$ then $\mathbf{f}(t) = e^{t\mathcal{L}}\mathbf{f}$ solves (1). For every $\mathbf{f} \in L^1_s(\mathcal{R})$ we have*

$$\|e^{t\mathcal{L}}\mathbf{f}\|_1 \le \|\mathbf{f}\|_1 .$$

Moreover, if $\mathbf{f}$ is positive then so is $e^{t\mathcal{L}}\mathbf{f}$ and $\|e^{t\mathcal{L}}\mathbf{f}\|_1 = \|\mathbf{f}\|_1$. Thus (1) generates an evolution that preserves probability distributions.

Proof See Sect. 3.1. □

It is not hard to see that the evolution generated by (1) admits the steady state $\boldsymbol{\Gamma}$ given by

$$(\boldsymbol{\Gamma})_N(\underline{v}_N) = \left(\frac{\mu}{\rho}\right)^N \frac{e^{-\frac{\mu}{\rho}}}{N!} e^{-\pi|\underline{v}_N|^2} := a_N \gamma_N(\underline{v}_N) \tag{7}$$

where $\gamma_N(\underline{v}_N) = \prod_{i=1}^N \gamma(v_i)$, with $\gamma(v) = e^{-\pi v^2}$, is the Maxwellian distribution with $\beta = 2\pi$ in dimension N while $a_N = \left(\frac{\mu}{\rho}\right)^N \frac{e^{-\frac{\mu}{\rho}}}{N!}$ is a Poisson distribution on $\mathbb{N}$. We observe that $\boldsymbol{\Gamma}$ is a Grand Canonical Ensemble with temperature $T = \beta^{-1} = 1/2\pi$, chemical potential $\chi = (2\pi)^{-1} \log(\rho/\mu)$, and average number of particles $\langle \mathcal{N}\boldsymbol{\Gamma}\rangle = \mu/\rho$ where

$$\langle \mathcal{N}\mathbf{f}\rangle := \sum_{N=0}^{\infty} N \int f_N(\underline{v}_N) d\underline{v}_N .$$

In Sect. 3.1 we show that $\boldsymbol{\Gamma}$ is the unique steady state of the evolution generated by (1). Finally, from a physical point of view, it is natural to consider only initial states with finite average number of particles and average kinetic energy, that is probability distributions $\mathbf{f}$ such that

$$\langle \mathcal{N}\mathbf{f}\rangle < \infty, \quad \text{and} \quad \langle \mathcal{E}\mathbf{f}\rangle := \sum_{N=0}^{\infty} \frac{1}{2} \int \Big(\sum_i v_i^2\Big) f_N(\underline{v}_N) d\underline{v}_N < \infty . \tag{8}$$

Since the Kac collision operator $\mathcal{K}$ preserves energy and number of particles, we can derive autonomous equations for the evolutions of $N(t) = \langle \mathcal{N}\mathbf{f}(t)\rangle$ and $E(t) = \langle \mathcal{E}\mathbf{f}(t)\rangle$. Indeed, if $\mathbf{f}$ is a probability distribution, we obtain

$$\begin{aligned} \frac{d}{dt}N(t) &= \mu - \rho N(t) \\ \frac{d}{dt}E(t) &= \frac{\mu}{2\pi} - \rho E(t) \end{aligned} \tag{9}$$

so that, if (8) holds at time $t = 0$ it holds for every time $t > 0$. See Sect. 3.1 for a derivation of these equations. Letting $e(t) = E(t)/N(t)$, we get

$$\frac{d}{dt}e(t) = \frac{\mu}{N(t)} \left(\frac{1}{2\pi} - e(t)\right) . \tag{10}$$

Equation (10) looks like Newton law of cooling for a system like ours. Notwithstanding this, $e(t)$ is not the natural definition of temperature since it is not the average kinetic energy per particle. A more interesting quantity is $\tilde{e}(t) = \langle v_1^2 \mathbf{f}\rangle$, but we were not able to obtain a closed form expression for its evolution.

As discussed in the introduction, we are interested in properties that are uniform in the average number of particles in the steady state $\langle \mathcal{N}\boldsymbol{\Gamma}\rangle = \mu/\rho$ and eventually we want to consider the situation where the average number of particles goes to infinity, that is $\mu/\rho \to \infty$. A classical way to take such a limit is to require that the collision rate between particles decreases as the average number of particles increases in such a way that the average number of collisions a given particle suffers in a given time is independent from μ/ρ, at least when μ/ρ is large. This is achieved by setting

$$\tilde{\lambda} = \lambda \frac{\rho}{\mu} .$$

Observe that in this way, the scaling in N of K_N in (5) differs from the scaling in the standard Kac model. Notwithstanding this, they can both be thought as implementations of the Grad-Boltzmann limit in the two different situations, see [11].

One way to study the approach of an initial state $\mathbf{f}$ toward $\boldsymbol{\Gamma}$ is by computing the spectral gap of $\mathcal{L}$. Since $\mathcal{L}$ is not self adjoint on $L^2_s(\mathcal{R})$ we perform a ground state transformation setting

$$f_N := a_N \gamma_N h_N . \tag{11}$$

We will express (11) as $\mathbf{f} = \boldsymbol{\Gamma}\mathbf{h}$. Inserting the above definition in (1) we get

$$\frac{d}{dt}\mathbf{h} = \widetilde{\mathcal{L}}\mathbf{h} := \rho(\mathcal{P}^+\mathbf{h} - \mathcal{N}\mathbf{h}) + \mu(\mathcal{P}^-\mathbf{h} - \mathbf{h}) + \tilde{\lambda}\mathcal{K}\mathbf{h}$$

where we have set

$$(\mathcal{P}^+\mathbf{h})_N = \sum_{i=1}^{N} h_{N-1}(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_N)$$

$$(\mathcal{P}^-\mathbf{h})_N = \frac{1}{N+1}\sum_{i=1}^{N+1}\int dw e^{-\pi w^2} h_{N+1}(v_1, \ldots, v_{i-1}, w, v_i, \ldots, v_N)$$

In this representation, the steady state is given by the vector $\mathbf{e}^0$ such that $(\mathbf{e}^0)_N \equiv 1$ for every N. Thus $\widetilde{\mathcal{L}}$ is an unbounded operator on the Hilbert space

$$L^2_s(\mathcal{R}, \boldsymbol{\Gamma}) = \bigoplus_{N=0}^{\infty} L^2_s(\mathbb{R}^N, a_N\gamma_N(\underline{v}_N))$$

of all states $\mathbf{h} = (h_0, h_1, h_2, \ldots)$ with $h_N(\underline{v}_N)$ symmetric under permutations of the v_i and defined by the scalar product

$$(\mathbf{h}_1, \mathbf{h}_2) := \sum_{N=0}^{\infty} a_N (h_{1,N}, h_{2,N})_N := \sum_{N=0}^{\infty} a_N \int h_{1,N}(\underline{v}_N) h_{2,N}(\underline{v}_N)\gamma_N(\underline{v}_N) d\underline{v}_N .$$

As for $\mathcal{L}$, defining $\widetilde{\mathcal{L}}_M\mathbf{h} = (\mathcal{L}\mathbf{h})_M$ we get a bounded operator from $L^2_s(\mathcal{R}, \boldsymbol{\Gamma})$ to $L^2_s(\mathbb{R}^N, \gamma_N(\underline{v}_N))$ so that, calling $\|h_N\|_{2,N} = (h_N, h_N)_N$, we can take

$$D^2 = \left\{\mathbf{h} \,\Big|\, \sum_{N=0}^{\infty} a_N \|(\widetilde{\mathcal{L}}\mathbf{h})_N\|_{2,N} < \infty\right\}$$

as the domain of $\widetilde{\mathcal{L}}$. The following Theorem shows that $\widetilde{\mathcal{L}}$ defines an evolution on $L^2_s(\mathcal{R}, \boldsymbol{\Gamma})$.

Theorem 2 *The generator $\widetilde{\mathcal{L}}$ is self adjoint and non-positive definite on $L^2_s(\mathcal{R}, \boldsymbol{\Gamma})$. Furthermore, if we define*

$$\Delta = \sup\{(\mathbf{h}, \widetilde{\mathcal{L}}\mathbf{h}) \mid \mathbf{h} \in D^2, \|\mathbf{h}\|_2 = 1, \mathbf{h} \perp \mathbf{E}_0\}$$

where $\|\mathbf{h}\|_2 = (\mathbf{h}, \mathbf{h})$ *and* $\mathbf{E}_0 = \text{span}\{\mathbf{e}^0\}$, *we get*

$$\Delta = -\rho\,.$$

Moreover Δ is an eigenvalue and the associated eigenspace is $\mathbf{E}_1 = \text{span}\{\mathbf{e}_1, \mathbf{e}_{(0,0,1)}\}$ *with* $\mathbf{e}_1 = \sqrt{\frac{\rho}{\mu}}\mathcal{P}^+\mathbf{e}^0 - \sqrt{\frac{\mu}{\rho}}\mathbf{e}^0$ *while*

$$(\mathbf{e}_{(0,0,1)})_N(\underline{v}_N) = \sqrt{\frac{\rho}{2\mu}}\sum_{i=1}^{N}(2\pi v_i^2 - 1)\,.$$

Proof See Sect. 3.2. □

Due to the invariance of even, second degree polynomials under the Kac collision operator $\mathcal{K}$, Theorem 2 shows that the spectral gap of the generator $\widetilde{\mathcal{L}}$ is completely determined by the presence of the reservoir. This is not surprising since all states $\mathbf{h}$ such that h_N is rotationally invariant for every N are in the null space of $\mathcal{K}$.

As in [4], to see the effect of the Kac collision operator $\mathcal{K}$, we have to look at the second gap, defined as

$$\Delta_2 = \sup\{(\mathbf{h}, \widetilde{\mathcal{L}}\mathbf{h}) \mid \mathbf{h} \in D^2, \|\mathbf{h}\|_2 = 1, \mathbf{h} \perp \mathbf{E}_0 \oplus \mathbf{E}_1\}\,. \tag{12}$$

Theorem 3 *If*

$$\rho > \frac{\lambda}{4} + 2\lambda\sqrt{\frac{\rho}{\mu}} \quad \text{and} \quad \frac{\mu}{\rho} > 256 \tag{13}$$

we have

$$-\rho - \frac{\lambda}{4} \le \Delta_2 < -\rho - \frac{\lambda}{4} + 2\lambda\sqrt{\frac{\rho}{\mu}}\,.$$

Moreover Δ_2 is an eigenvalue and the associated eigenspace is contained in the space of all states $\mathbf{h}$ *such that h_N is an even, fourth degree polynomial.*

Proof See Sect. 3.3. □

Since μ/ρ is the average number of particles in the steady state, the conditions in (13) are not too restrictive.

It is possible to see that, as in the case of the standard Kac evolution, the L^2 norm discussed above does not scale well with the average number of particles in the system and thus it is not a good measure of distance from the steady state if μ/ρ is large. A better measure is the entropy of a probability distribution $\mathbf{f}$ relative to the steady state $\boldsymbol{\Gamma}$ defined as

$$\mathcal{S}(\mathbf{f} \mid \boldsymbol{\Gamma}) = \sum_N a_N \int dv_N h_N(\underline{v}_N)\log h_N(\underline{v}_N)\gamma_N(\underline{v}_N)$$

where, as before, $\mathbf{f} = \boldsymbol{\Gamma}\mathbf{h}$ and a_N and γ_N are defined in (11).

As usual, it is easy to show using convexity that $\mathcal{S}(\mathbf{f} \mid \boldsymbol{\Gamma}) \ge 0$, $\mathcal{S}(\mathbf{f} \mid \boldsymbol{\Gamma}) = 0$ if and only if $\mathbf{f} = \boldsymbol{\Gamma}$. Moreover, from Lemma 1 and convexity, it follows that $\mathcal{S}(\mathbf{f}(t) \mid \boldsymbol{\Gamma}) \le \mathcal{S}(\mathbf{f} \mid \boldsymbol{\Gamma})$ where $\mathbf{f}(t) = e^{t\mathcal{L}}\mathbf{f}$. In Sect. 3.4, we show that, thanks to the presence of the reservoir, the entropy production rate is strictly negative. More precisely, assuming that $\mathbf{f} = \boldsymbol{\Gamma}\mathbf{h} \in D^1$ and $\boldsymbol{\Gamma}\mathbf{h}\log\mathbf{h} \in D^1$ we essentially obtain that

$$\frac{d}{dt}\mathcal{S}(\mathbf{f}(t)\,|\,\boldsymbol{\Gamma}) \le -\rho\mathcal{S}(\mathbf{f}(t)\,|\,\boldsymbol{\Gamma})\,. \tag{14}$$

See Lemmas 19 and 20 in Sect. 3.4 below for a precise statement. Form (14) we obtain the following Theorem.

Theorem 4 *If $\mathbf{f} = \mathbf{h}\boldsymbol{\Gamma} \in D^1$ is a probability distribution such that $\boldsymbol{\Gamma}\mathbf{h}\log\mathbf{h} \in D^1$ then*

$$\mathcal{S}(\mathbf{f}(t)\,|\,\boldsymbol{\Gamma}) \le e^{-\rho t}\mathcal{S}(\mathbf{f}(0)\,|\,\boldsymbol{\Gamma})\,. \tag{15}$$

Proof See Sect. 3.4. □

As in the case of Theorem 2, convergence to equilibrium in entropy is completely dominated by the presence of the thermostat, that is, Theorem 4 remains valid in the case $\tilde{\lambda} = 0$ where there is no collision among the particles.

We can now discuss the validity of a Boltzmann-Kac type equation when the average number of particles in the system goes to infinity. To follow the standard analysis in [16], we have first to define what a *chaotic sequence* is in the present situation. It is natural to call $\mathbf{f} = (f_0, f_1, f_2, \ldots)$ a *product state* if it has the form

$$f_N(\underline{v}_N) = e^{-\eta}\frac{\eta^N}{N!}\prod_{i=1}^N g(v_i) \tag{16}$$

where $g(v)$ is a probability density on $\mathbb{R}$ and $\eta > 0$ is the average number of particles. We observe that for the state $\mathbf{f}$ in (16), we have

$$\left(e^{t\mathcal{T}}\mathbf{f}\right)_N = e^{-\eta(t)}\frac{\eta(t)^N}{N!}\prod_{i=1}^N g(v_i, t) \tag{17}$$

where $\mathcal{T}$ is defined in (4) and, calling $l(v,t) = \frac{\rho}{\mu}\eta(t)g(v,t)$, we get

$$\begin{aligned} \eta(t) &= e^{-\rho t}\eta + (1 - e^{-\rho t})\frac{\mu}{\rho} \\ l(v,t) &= e^{-\rho t}l(v) + (1 - e^{-\rho t})\gamma(v) \end{aligned} \tag{18}$$

This implies that the thermostat preserves the product structure exactly. See Sect. 3.5 for a derivation of (17) and (18).

Thus we call a sequence of states $\mathbf{f}_n = (f_{n,0}, f_{n,1}, f_{n,2}, \ldots)$ chaotic if it approaches the structure (16) while the average number of particles $\langle\mathcal{N}\mathbf{f}_n\rangle$ goes to infinity. More precisely, let μ_n be a sequence such that $\lim_{n\to\infty}\mu_n = \infty$ and define

$$F_n^{(k)}(\underline{v}_k) = \left(\frac{\rho}{\mu_n}\right)^k \sum_{N\ge k}\frac{N!}{(N-k)!}\int f_{n,N}(\underline{v}_k, \underline{v}_{N-k})d\underline{v}_{N-k} \tag{19}$$

where the factor $\frac{N!}{(N-k)!}$ accounts for the possible ways to choose the k particles with velocities $\underline{v}_k$. We also define

$$\|\mathbf{f}\|_1^{(k)} = \sum_{N\ge k}\frac{N!}{(N-k)!}\|f_N\|_{1,N} \tag{20}$$

so that $\|F_n^{(k)}\|_{1,k} \le \left(\frac{\rho}{\mu_n}\right)^k \|\mathbf{f}_n\|_1^{(k)}$.

Observe that, if $\mathbf{f}_n$ is a product state of the form (16) with average number of particles η_n, that is if

$$f_{n,N}(\underline{v}_N) = e^{-\eta_n} \frac{\eta_n^N}{N!} \prod_{i=1}^{N} g(v_i)$$

we get

$$F_n^{(k)}(\underline{v}_k) = \left(\frac{\eta_n \rho}{\mu_n} \right)^k \prod_{i=1}^{k} g(v_i) \, .$$

Thus the factor $\left(\frac{\rho}{\mu_n}\right)^k$ in (19) assures that, at least in this case, if $\lim_{n\to\infty} \eta_n/\mu_n$ exists then also $\lim_{n\to\infty} F_n^{(k)}$ exists.

To generalize these observations, we say that $F_n^{(k)}$ converges weakly to $F^{(k)}$ if, for any continuous and bounded test function $\phi_k : \mathbb{R}^k \to \mathbb{R}$, we have

$$\lim_{n\to\infty} \int_{\mathbb{R}^k} F_n^{(k)}(\underline{v}_k)\phi_k(\underline{v}_k) d\underline{v}_k = \int_{\mathbb{R}^k} F^{(k)}(\underline{v}_k)\phi_k(\underline{v}_k) d\underline{v}_k$$

and we write $\text{w-lim}_{n\to\infty} F_n^{(k)} = F^{(k)}$. Given a sequence $\mathbf{f}_n$ of probability distributions such that

$$\|\mathbf{f}_n\|_1^{(r)} \leq M^r \left(\frac{\mu_n}{\rho} \right)^r \tag{21}$$

for some $M > 0$ and every n and r, we say that $\mathbf{f}_n$ is *chaotic (w.r.t.* μ_n*)* if, for some F

$$\underset{n\to\infty}{\text{w-lim}}\, F_n^{(1)} = F \tag{22}$$

while for every $k > 1$ we have

$$\underset{n\to\infty}{\text{w-lim}}\, F_n^{(k)} = F^{\otimes k} \tag{23}$$

where $F^{\otimes k}(\underline{v}_k) = \prod_{i=1}^{k} F(v_i)$. Observe that

$$\int F(v)dv = \lim_{n\to\infty} \frac{\langle \mathcal{N} \mathbf{f}_n \rangle \rho}{\mu_n} \tag{24}$$

so that we can see $F(v)$ as the *relative particle density*.

In [14, 16] a sequence of probability distributions $f_n : \mathbb{R}^n \to \mathbb{R}$ is said to be chaotic if, calling

$$\widetilde{F}_n^{(k)}(\underline{v}_k) = \int f_n(\underline{v}_k, \underline{v}_{n-k}) d\underline{v}_{n-k} \, ,$$

we have

$$\underset{n\to\infty}{\text{w-lim}}\, \widetilde{F}_n^{(1)} = \widetilde{F} \qquad \text{and} \qquad \underset{n\to\infty}{\text{w-lim}}\, \widetilde{F}_n^{(k)} = \widetilde{F}^{\otimes k} \, .$$

If we consider the sequence of states $\mathbf{f}_n$ defined as

$$(\mathbf{f}_n)_N = \begin{cases} f_n & n = N \\ 0 & n \neq N \end{cases}$$

with the natural choice $\mu_n = n\rho$, since the number of particles in $\mathbf{f}_n$ is exactly n, from (19) we get $F = \widetilde{F}$ and thus $F^{(k)} = \widetilde{F}^{(k)}$. In this sense, (19) and (23) can be considered as a generalization of the classical definition in [14].

Let now

$$\mathbf{f}_n(t) = e^{\mathcal{L}_n t}\mathbf{f}_n(0)$$

where $\mathcal{L}_n$ is given by (1) with $\mu = \mu_n$ and

$$\tilde{\lambda} = \tilde{\lambda}_n = \lambda \frac{\rho}{\mu_n}\,. \tag{25}$$

In Sect. 3.6, we prove that $e^{\mathcal{L}_n t}$ propagates chaos in the sense that, if $\mathbf{f}_n(0)$ forms a chaotic sequence, then $\mathbf{f}_n(t)$ also forms a chaotic sequence for every t. This gives the following theorem.

Theorem 5 *If $\mathbf{f}_n(0)$ forms a chaotic sequence w.r.t. μ_n, with $\lim_{n\to\infty}\mu_n = \infty$, then also $\mathbf{f}_n(t)$ forms a chaotic sequence for every $t \geq 0$. Moreover the relative particle density*

$$F(v,t) = \underset{n\to\infty}{w\text{-}lim}\,\frac{\rho}{\mu_n}\sum_{N=1}^{\infty} N\int f_{n,N}(v,\underline{v}_{N-1},t)d\underline{v}_{N-1}$$

satisfies the Boltzmann-Kac type equation

$$\begin{aligned}\frac{d}{dt}F(v,t) = &-\rho(F(v,t)-\gamma(v))\\ &+\lambda\int_{\mathbb{R}}dw\int\frac{d\theta}{2\pi}[F(v\cos\theta + w\sin\theta,t)\\ &\times F(-v\sin\theta + w\cos\theta,t) - F(w,t)F(v,t)]\,.\end{aligned} \tag{26}$$

Proof See Sect. 3.6. □

3 Proofs

3.1 Proof of Lemma 1

The results in this section are based on two observations. The first is that the collision operator $\mathcal{K}$ acts independently on each $L^1_s(\mathbb{R}^N)$ and thus preserves positivity and probability. The second is that, due to the different scaling in N of the *in* and *out* operators, see (2) and (3), for large N the outflow of particles dominates the inflow. Thus even if the initial probability of having a number of particles much larger than the steady state average μ/ρ is high, this probability will rapidly decrease toward its steady state value, see (34) and (46) below. In particular this prevents probability from "leaking out at infinity".

We will now construct a solution of (1) in three steps, starting from $\mathcal{K}$ alone, using a partial power series expansion, see (27) below, and then adding the *out* operator $\mathcal{O}$ and finally the *in* operator $\mathcal{I}$, using a Duhamel style expansion, see (33) and (40) below. These expansions are strongly inspired by the stochastic nature of the the evolution studied, see Remark 8 below for more details.

It is natural to define $\left(e^{t\tilde{\lambda}\mathcal{K}}\mathbf{f}\right)_N = e^{t\tilde{\lambda}K_N}f_N$ where we can write

$$e^{t\tilde{\lambda}K_N}f_N = e^{-\tilde{\lambda}t\binom{N}{2}}\sum_{n=0}^{\infty}\frac{\tilde{\lambda}^n t^n Q_N^n}{n!}f_N\,. \tag{27}$$

Observing that

$$\begin{aligned}\|e^{t\tilde{\lambda}K_N} f_N - f_N\|_{1,N} &\leq \left(1 - e^{-\tilde{\lambda}t\binom{N}{2}}\right)\|f_N\|_{1,N} + \left\|e^{-\tilde{\lambda}t\binom{N}{2}}\sum_{n=1}^{\infty}\frac{\tilde{\lambda}^n t^n Q_N^n}{n!}f_N\right\|_{1,N}\\ &\leq 2\left(1 - e^{-\tilde{\lambda}t\binom{N}{2}}\right)\|f_N\|_{1,N} \end{aligned}\tag{28}$$

and using that from Dominated Convergence we get

$$\lim_{t\to 0^+}\sum_{N=0}^{\infty}\left(1 - e^{-\tilde{\lambda}t\binom{N}{2}}\right)\|f_N\|_{1,N} = 0$$

we obtain that $\lim_{t\to 0^+} e^{t\tilde{\lambda}\mathcal{K}}\mathbf{f} = \mathbf{f}$. Similarly, we get

$$\begin{aligned}&\frac{1}{t}\|e^{t\tilde{\lambda}K_N} f_N - f_N - \tilde{\lambda}tK_N f_N\|_{1,N}\\ &\quad\leq \frac{1}{t}\left(e^{-\tilde{\lambda}t\binom{N}{2}} - 1 + \tilde{\lambda}t\binom{N}{2}\right)\|f_N\|_{1,N} + \left(1 - e^{-\tilde{\lambda}t\binom{N}{2}}\right)\|\tilde{\lambda}Q_N f_N\|_{1,N}\\ &\qquad + \frac{1}{t}\left\|e^{-\tilde{\lambda}t\binom{N}{2}}\sum_{n=2}^{\infty}\frac{\tilde{\lambda}^n t^n Q_N^n}{n!}f_N\right\|_{1,N}\\ &\quad\leq \frac{2}{t}\left(e^{-\tilde{\lambda}t\binom{N}{2}} - 1 + \tilde{\lambda}t\binom{N}{2}\right)\|f_N\|_{1,N} + \tilde{\lambda}\binom{N}{2}\left(1 - e^{-\tilde{\lambda}t\binom{N}{2}}\right)\|f_N\|_{1,N}\end{aligned}$$

so that, if $\mathbf{f} \in D_1$ then $\lim_{t\to 0^+}\left(e^{t\tilde{\lambda}\mathcal{K}}\mathbf{f} - \mathbf{f}\right)/t = \tilde{\lambda}\mathcal{K}\mathbf{f}$. Since $\|e^{t\tilde{\lambda}K_N} f_N\|_{1,N} \leq \|f_N\|_{1,N}$ we get $\|e^{t\tilde{\lambda}\mathcal{K}}\mathbf{f}\|_1 \leq \|\mathbf{f}\|_1$. Moreover if f_N is positive then also $e^{t\tilde{\lambda}K_N} f_N$ is positive and $\|e^{t\tilde{\lambda}K_N} f_N\|_{1,N} = \|f_N\|_{1,N}$. Thus if $\mathbf{f}$ is positive then $e^{t\tilde{\lambda}\mathcal{K}}\mathbf{f}$ is positive and $\|e^{t\tilde{\lambda}\mathcal{K}}\mathbf{f}\|_1 = \|\mathbf{f}\|_1$.

Let now $\mathbf{f}(t)$ be a solution of

$$\frac{d}{dt}\mathbf{f}(t) = \tilde{\lambda}\mathcal{K}\mathbf{f}(t) + \rho(\mathcal{O} - \mathcal{N})\mathbf{f}(t)\tag{29}$$

with $\mathbf{f}(0) = \mathbf{f} \in D^1$. If such a solution exists, it satisfies the Duhamel formula

$$f_N(t) = e^{(\tilde{\lambda}K_N - \rho N)t} f_N + \rho\int_0^t e^{(\tilde{\lambda}K_N - \rho N)(t-s)}\left(\mathcal{O}\mathbf{f}(s)\right)_N\, ds\tag{30}$$

where the construction of $e^{(\tilde{\lambda}\mathcal{K} - \rho\mathcal{N})t}$ is analogous to that of $e^{\tilde{\lambda}\mathcal{K}t}$. From (30) we get

$$\|f_N(t)\|_{1,N} \leq e^{-\rho Nt}\|f_N\|_{1,N} + \int_0^t e^{-\rho N(t-s)}\rho(N+1)\,\|f_{N+1}(s)\|_{1,N+1}\, ds\tag{31}$$

where we have used that

$$\|(\mathcal{O}\mathbf{f})_N\|_{1,N} = (N+1)\int\left|\int f_{N+1}(\underline{v}_{N+1})dv_{N+1}\right| d\underline{v}_N \leq (N+1)\|f_{N+1}\|_{1,N+1}\,.\tag{32}$$

Observe that, in (32), equality holds if and only if f_{N+1} is everywhere positive or everywhere negative. To construct a solution of (29) we iterate (30) to define

$$\begin{aligned}\mathcal{Q}(t)\mathbf{f} = e^{(\tilde{\lambda}\mathcal{K} - \rho\mathcal{N})t}\mathbf{f} + \sum_{n=1}^{\infty}\int_{0<t_1<\ldots<t_n<t} & e^{(\tilde{\lambda}\mathcal{K} - \rho\mathcal{N})(t-t_n)}\rho\mathcal{O}e^{(\tilde{\lambda}\mathcal{K} - \rho\mathcal{N})(t_n - t_{n-1})}\\ &\cdots\rho\mathcal{O}e^{(\tilde{\lambda}\mathcal{K} - \rho\mathcal{N})t_1}\mathbf{f}\, dt_1\cdots dt_n\end{aligned}\tag{33}$$

and then show that $\mathcal{Q}(t)$ is a semigroup of bounded operators and that $\mathbf{f}(t) = \mathcal{Q}(t)\mathbf{f}$ solves (29) if $\mathbf{f} \in D^1$. Using (31) iteratively we get

$$\begin{aligned}\left\|(\mathcal{Q}(t)\mathbf{f})_N\right\|_{1,N} &\leq \sum_{n\geq 0} e^{-\rho N t}\frac{(N+n)!}{N!}\int_{0<t_1<\cdots<t_n<t} \\ &\quad \prod_{i=1}^{n} e^{\rho(N+n-i)t_i}\rho e^{-\rho(N+n-i+1)t_i}dt_1\cdots dt_n\|f_{N+n}\|_{1,N+n} \\ &= e^{-\rho N t}\sum_{n\geq 0}\binom{N+n}{N}\left(1-e^{-\rho t}\right)^n\|f_{N+n}\|_{1,N+n} \end{aligned} \tag{34}$$

where, in the last identity, we have used that

$$\rho^n\int_{0\leq t_1\leq\cdots t_n\leq t}\prod_{i=1}^{n} e^{-\rho t_i}dt_1\cdots dt_n = \frac{1}{n!}(1-e^{-\rho t})^n\,. \tag{35}$$

After summing over N we get

$$\|\mathcal{Q}(t)\mathbf{f}\|_1 \leq \sum_{N\geq 0}\sum_{n\geq 0}\binom{N+n}{N}e^{-\rho N t}\left(1-e^{-\rho t}\right)^n\|f_{N+n}\|_{1,N+n} = \|\mathbf{f}\|_1\,. \tag{36}$$

so that $\|\mathcal{Q}(t)\|_1 \leq 1$. Observe also that, if $\mathbf{f}$ is positive then $\mathcal{Q}(t)\mathbf{f}$ is positive and $\|\mathcal{Q}(t)\mathbf{f}\|_1 = \|\mathbf{f}\|_1$, see comment below (32). Conversely, if for some N, f_N takes both positive and negative values then $\|\mathcal{Q}(t)\mathbf{f}\|_1 < \|\mathbf{f}\|_1$.

From (33), we see that $\mathcal{Q}(t_1)\mathcal{Q}(t_2) = \mathcal{Q}(t_1+t_2)$ while, using (34) and (36), and the fact that

$$N(N-1)\binom{M}{N} = M(M-1)\binom{M-2}{N-2}$$

we get

$$\sum_{N=1}^{\infty} N^2\left\|(\mathcal{Q}(t)\mathbf{f})_N\right\|_{1,N} \leq e^{-\rho t}\sum_{N=1}^{\infty}N^2\|f_N\|_{1,N}$$

so that $\mathcal{Q}(t)\mathbf{f} \in D^1$ if $\mathbf{f} \in D^1$. Moreover observe that

$$\begin{aligned}\|\mathcal{Q}(t)\mathbf{f}-\mathbf{f}\|_1 &\leq \sum_{N\geq 0}\sum_{n\geq 1}\binom{N+n}{N}e^{-\rho N t}\left(1-e^{-\rho t}\right)^n\|f_{N+n}\|_{1,N+n} \\ &\quad + \left\|e^{(\tilde{\lambda}\mathcal{K}-\rho\mathcal{N})t}\mathbf{f}-\mathbf{f}\right\|_1 \\ &= \sum_{N\geq 0}\left(1-e^{-\rho N t}\right)\|f_N\|_{1,N} + \left\|e^{(\tilde{\lambda}\mathcal{K}-\rho\mathcal{N})t}\mathbf{f}-\mathbf{f}\right\|_1 \end{aligned} \tag{37}$$

so that $\lim_{t\to 0^+}\mathcal{Q}(t)\mathbf{f} = \mathbf{f}$. Similarly we have

$$\begin{aligned}&\frac{1}{t}\|\mathcal{Q}(t)\mathbf{f}-\mathbf{f}-t(\tilde{\lambda}\mathcal{K}-\rho(\mathcal{O}-\mathcal{N}))\mathbf{f}\|_1 \\ &\quad\leq \frac{1}{t}\sum_{N\geq 0}\sum_{n\geq 2}\binom{N+n}{N}e^{-\rho N t}\left(1-e^{-\rho t}\right)^n\|f_{N+n}\|_{1,N+n} \\ &\qquad + \frac{1}{t}\left\|e^{(\tilde{\lambda}\mathcal{K}-\rho\mathcal{N})t}-\mathbf{f}-t(\tilde{\lambda}\mathcal{K}-\rho\mathcal{N})\mathbf{f}\right\|_1 \end{aligned}$$

$$+\frac{\rho}{t}\left\|\int_0^t e^{(\tilde{\lambda}\mathcal{K}-\rho\mathcal{N})(t-s)}\mathcal{O}e^{(\tilde{\lambda}\mathcal{K}-\rho\mathcal{N})s}\mathbf{f}-t\mathcal{O}\mathbf{f}\right\|_1 . \tag{38}$$

If $\mathbf{f}\in D^1$, proceeding as in (37) we see that the second and third lines of the right hand side of (38) vanish as $t\to 0^+$ while writing

$$\int_0^t e^{(\tilde{\lambda}\mathcal{K}-\rho\mathcal{N})(t-s)}\mathcal{O}e^{(\tilde{\lambda}\mathcal{K}-\rho\mathcal{N})s}\mathbf{f}-t\mathcal{O}\mathbf{f}=\int_0^t e^{(\tilde{\lambda}\mathcal{K}-\rho\mathcal{N})(t-s)}\mathcal{O}\left(e^{(\tilde{\lambda}\mathcal{K}-\rho\mathcal{N})s}\mathbf{f}-\mathbf{f}\right)$$
$$+\int_0^t\left(e^{(\tilde{\lambda}\mathcal{K}-\rho\mathcal{N})(t-s)}\mathcal{O}-\mathcal{O}\right)\mathbf{f} \tag{39}$$

and using (28) we see that also the last line of (38) vanish as $t\to 0^+$. This implies that, for $\mathbf{f}\in D^1$, we have $\lim_{t\to 0^+}(\mathcal{Q}(t)\mathbf{f}-\mathbf{f})/t=\tilde{\lambda}\mathcal{K}\mathbf{f}+\rho(\mathcal{O}-\mathcal{N})\mathbf{f}$ and we can write $\mathcal{Q}(t)=e^{t(\tilde{\lambda}\mathcal{K}+\rho(\mathcal{O}-\mathcal{N}))}$.

We can now use a Duhamel style expansion once more to obtain

$$e^{t\mathcal{L}}\mathbf{f}=e^{(\tilde{\lambda}\mathcal{K}+\rho(\mathcal{O}-\mathcal{N})-\mu\mathrm{Id})t}\mathbf{f}$$
$$+\sum_{n=1}^{\infty}\mu^n\int\limits_{0<t_1<\ldots<t_n<t} e^{(\tilde{\lambda}\mathcal{K}+\rho(\mathcal{O}-\mathcal{N})-\mu\mathrm{Id})(t-t_n)}\mathcal{I}e^{(\tilde{\lambda}\mathcal{K}+\rho(\mathcal{O}-\mathcal{N})-\mu\mathrm{Id})(t_n-t_{n-1})}$$
$$\cdots\mathcal{I}e^{(\tilde{\lambda}\mathcal{K}+\rho(\mathcal{O}-\mathcal{N})-\mu\mathrm{Id})t_1}\mathbf{f}\,dt_1\cdots dt_n \tag{40}$$

that, thanks to the fact that $\mathcal{I}$ is bounded, converges for every $\mathbf{f}\in L^1(\mathcal{R})$ to a solution of $\frac{d}{dt}\mathbf{f}(t)=\mathcal{L}\mathbf{f}(t)$. Lemma 1 follows easily observing that $\|\mathcal{I}\mathbf{f}\|_1=\|\mathbf{f}\|_1$. □

Remark 6 The proof of Lemma 1 above also shows that given $\mathbf{f}\in L^1(\mathcal{R})$, if for some N, f_N takes both positive and negative values, then $\|e^{t\mathcal{L}}\mathbf{f}\|_1<\|\mathbf{f}\|_1$.

Remark 7 From (30) it is not hard to see that, if $\mathbf{f}_i(t)\in D^1$, $i=1,2$, are two solutions of (1) with $\mathbf{f}_1(0)=\mathbf{f}_2(0)$ then $\mathbf{f}_1(t)=\mathbf{f}_2(t)$.

Remark 8 Observe that (1) is the master equation of a jump process where jumps occur when two particles collide, a particle enters the system or a particle leaves it. Moreover, these jumps arrive according to a Poisson process. The expansions (27), (33) and (40) combined can be seen as a representation of the evolution of $\mathbf{f}$ as an integral over all possible realizations of the jump process, sometime called *jump* or *collision histories*. A similar representation was used in [2] to study the interaction of a Kac system with a large reservoir. Clearly, such a representation is much more complex in the present situation then for the model studied in [2]. Here the arrival rate for the jumps depends on the state of the system via the number of particles N and goes to infinity as N increases.

Given a state $\mathbf{f}=(f_0,f_1,\ldots)$ we set $\bar{f}_N=\int f_N(\underline{v}_N)d\underline{v}_N$. It is easy to see that

$$\int(\mathcal{O}\mathbf{f})_N(\underline{v}_N)d\underline{v}_N=(N+1)\bar{f}_{N+1}\,,\qquad\int(\mathcal{I}\mathbf{f})_N(\underline{v}_N)d\underline{v}_N=\bar{f}_{N-1} \tag{41}$$

while

$$\int(\mathcal{K}\mathbf{f})_N(\underline{v}_N)d\underline{v}_N=0\,,$$

so that we get

$$\overline{(\mathcal{L}\mathbf{f})_0}=-\mu\bar{f}_0+\rho\bar{f}_1$$

$$\overline{(\mathcal{L}\mathbf{f})_N} = -(N\rho + \mu)\bar{f}_N + \mu \bar{f}_{N-1} + \rho(N+1)\bar{f}_{N+1} \quad N > 0\,. \tag{42}$$

If $\boldsymbol{\Gamma}$ is a steady state, writing

$$\overline{\Gamma}_N = c_N \left(\frac{\mu}{\rho}\right)^N \frac{1}{N!}$$

we see from (42) that $c_N = c_0$ for every N. Since $\sum_N \overline{\Gamma}_N = 1$ we get $\overline{\Gamma}_N = a_N$, see (7). This implies that if $\boldsymbol{\Gamma}$ and $\boldsymbol{\Gamma}'$ are two steady states then

$$\int (\Gamma_N(\underline{v}_N) - \Gamma'_N(\underline{v}_N))d\underline{v}_N = 0$$

for every N. From Remark 6 it follows that, if $\boldsymbol{\Gamma} \neq \boldsymbol{\Gamma}'$ then $\|e^{t\mathcal{L}}(\boldsymbol{\Gamma} - \boldsymbol{\Gamma}')\|_1 < \|\boldsymbol{\Gamma} - \boldsymbol{\Gamma}'\|_1$. Uniqueness of the steady state follows immediately.

We now prove a more general version of (9). For $r \geq 0$ we define

$$N_r(\mathbf{f}) = \sum_{N=r}^{\infty} \frac{N!}{(N-r)!}\bar{f}_N\,. \tag{43}$$

and, using (42), we get

$$\begin{aligned} \frac{d}{dt}N_r(\mathbf{f}) &= \sum_{N=r}^{\infty} \frac{N!}{(N-r)!}\left(-(N\rho + \mu)\bar{f}_N + \mu \bar{f}_{N-1} + \rho(N+1)\bar{f}_{N+1}\right) \\ &= -\rho r N_r(\mathbf{f}) + \mu r N_{r-1}(\mathbf{f}) \end{aligned} \tag{44}$$

that, for $r = 1$, would imply the first of (9) since for a probability distribution we have $N_0(\mathbf{f}) = 1$. This argument is suggestive but only formal since we need to show that we can exchange the sum with the derivative in the above derivation. Notwithstanding this, it shows that for $r = 0$, if $\mathbf{f} \in D^1$ then

$$\sum_{N=0}^{\infty} \overline{(\mathcal{L}\mathbf{f})_N} = 0\,. \tag{45}$$

To prove (9) we proceed more directly using the expansions derived previously. Indeed from (34) and (36) we get

$$\begin{aligned} N_r\left(e^{t(\tilde{\lambda}\mathcal{K} - \rho\mathcal{N} + \rho\mathcal{O})}\mathbf{f}\right) &= \sum_{N \geq r}\sum_{n \geq 0} \frac{N!}{(N-r)!}\binom{N+n}{N} e^{-\rho N t}\left(1 - e^{-\rho t}\right)^n \bar{f}_{N+n} \\ &= e^{-\rho r t} N_r(\mathbf{f})\ . \end{aligned} \tag{46}$$

Furthermore, using that $N_r(\mathcal{I}\mathbf{f}) = N_r(\mathbf{f}) + r N_{r-1}(\mathbf{f})$, we get

$$\begin{aligned} N_r\left(\mathbf{f}(t)\right) &= N_r\left(e^{t(\tilde{\lambda}\mathcal{K} - \rho\mathcal{N} + \rho\mathcal{O} - \mu\mathrm{Id})}\mathbf{f}(0) + \mu\int_0^t e^{(t-s)(\tilde{\lambda}\mathcal{K} - \rho\mathcal{N} + \rho\mathcal{O} - \mu\mathrm{Id})}\mathcal{I}\mathbf{f}(s)ds\right) \\ &= e^{-(\rho r + \mu)t} N_r(\mathbf{f}(0)) + \mu\int_0^t e^{-(\rho r + \mu)(t-s)} N_r(\mathbf{f}(s))ds \\ &\quad + r\mu\int_0^t e^{-(\rho r + \mu)(t-s)} N_{r-1}(\mathbf{f}(s))ds \end{aligned} \tag{47}$$

that gives

$$N_r\left(\mathbf{f}(t)\right) = e^{-\rho r t} N_r(\mathbf{f}(0)) + r\mu\int_0^t e^{-\rho r(t-s)} N_{r-1}(\mathbf{f}(s))ds\,. \tag{48}$$

For $r = 1$, if $\mathbf{f}(0)$ is a probability distribution, we get

$$N(t) = e^{-\rho t} N(0) + (1 - e^{-\rho t}) \frac{\mu}{\rho}$$

that proves the first of (9). We will need the following corollary in Sect. 3.6 below.

Corollary 9 *Given a probability distribution* $\mathbf{f}$*, assume that there exists* M *such that* $|N_r(\mathbf{f}(0))| \leq M^r$ *then we have*

$$|N_r(\mathbf{f}(t))| \leq \max \left\{ M, \frac{\mu}{\rho} \right\}^r \tag{49}$$

for every $t \geq 0$.

Proof Clearly (49) holds for $r = 0$ since $N_0(\mathbf{f}(t)) = 1$ for every $t \geq 0$. Calling $M_1 = \max \left\{ M, \frac{\mu}{\rho} \right\}$, assume that $|N_{r-1}(\mathbf{f}(t))| \leq M_1^{r-1}$. Form (48) we get

$$\begin{aligned} |N_r(\mathbf{f}(t))| &\leq e^{-\rho r t} M^r + r\mu \int_0^t e^{-\rho r(t-s)} M_1^{r-1} ds \\ &= e^{-\rho r t} M^r + \frac{\mu}{\rho}(1 - e^{-\rho r t}) M_1^{r-1} \leq \max \left\{ M^r, \frac{\mu}{\rho} M_1^{r-1} \right\}. \end{aligned}$$

The corollary follows by induction on r. □

Let now

$$\tilde{f}_N = \sum_{i=1}^N \int v_i^2 f_N(\underline{v}_N) d\underline{v}_N$$

so that $E(t) = \sum_{N=1}^\infty \tilde{f}_N$ and observe that

$$\begin{aligned} \sum_{i=1}^N \int v_i^2 (\mathcal{O}\mathbf{f})_N(\underline{v}_N) d\underline{v}_N &= N \tilde{f}_{N+1}. \\ \sum_{i=1}^N \int v_i^2 (\mathcal{I}\mathbf{f})_N(\underline{v}_N) d\underline{v}_N &= \tilde{f}_{N-1} + \frac{1}{2\pi} \bar{f}_{N-1} \end{aligned}$$

while

$$\sum_{i=1}^N \int v_i^2 (\mathcal{K}\mathbf{f})_N(\underline{v}_N) d\underline{v}_N = 0.$$

Again proceeding formally we get

$$\begin{aligned} \frac{d}{dt} \sum_{N=1}^\infty \tilde{f}_N &= \sum_{N=1}^\infty \left(-(N\rho + \mu) \tilde{f}_N + \mu \tilde{f}_{N-1} + \frac{\mu}{2\pi} \bar{f}_{N-1} + \rho N \tilde{f}_{N+1} \right) \\ &= \frac{\mu}{2\pi} \sum_{N=0}^\infty \bar{f}_N - \rho \sum_{N=1}^\infty \tilde{f}_N. \end{aligned}$$

It is not hard to adapt this argument, together with (46) and (47), to prove the second of (9).

3.2 Proof of Theorem 2

To prove Theorems 2 and 3, we will construct a basis of eigenvectors for the generator

$$\mathcal{G} = \rho(\mathcal{P}^+ - \mathcal{N}) + \mu(\mathcal{P}^- - \mathrm{Id})$$

of the evolution due to the thermostat on $L^2_s(\mathcal{R}, \boldsymbol{\Gamma})$. We start by defining

$$(\mathcal{P}^+(g)\mathbf{h})_N(\underline{v}_N) = \sum_{i=1}^{N} h_{N-1}(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_N)g(v_i)$$
$$(\mathcal{P}^-(g)\mathbf{h})_N(\underline{v}_N) = \frac{1}{N+1}\sum_{i=1}^{N+1}\int dw e^{-\pi w^2} g(w) h_{N+1}(\underline{v}_{N,i}(w)) \tag{50}$$

with $\underline{v}_{N,i}(w) = (v_1, \ldots, v_{i-1}, w, v_i, \ldots, v_N)$ and $g \in L^2(\mathbb{R}, \gamma)$. Moreover, we use the convention that the sum over an empty set is 0 so that $(\mathcal{P}^+(g)\mathbf{h})_0 = 0$ for every $\mathbf{h}$. With this notation, $\mathcal{P}^+$ and $\mathcal{P}^-$ from the introduction are $\mathcal{P}^+(1)$ and $\mathcal{P}^-(1)$, respectively.

Lemma 10 *We have*

$$\rho\mathcal{P}^+(g)^* = \mu\mathcal{P}^-(g) \tag{51}$$

so that $\mathcal{G}$ is self-adjoint.

Proof Proceeding as in the definition of D^2, we take as domain of $\mathcal{P}^\pm(g)$ the subspaces

$$D^\pm = \left\{\mathbf{h}\,\middle|\, \sum_{N=0}^{\infty} a_N \|(\mathcal{P}^\pm(g)\mathbf{h})_N\|^2_{2,N} < \infty\right\}.$$

It is easy to see that $D^\pm$ are dense in $L^2(\mathcal{R}, \boldsymbol{\Gamma})$.

Calling $\underline{v}^i_N = (v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_N)$ we get

$$\begin{aligned}
(h_N, (\mathcal{P}^+(g)\mathbf{j})_N)_N &= \sum_{i=1}^{N}\int d\underline{v}_N \gamma_N(\underline{v}_N) h_N(\underline{v}_N) j_{N-1}(\underline{v}^i_N) g(v_i)\\
&= \sum_{i=1}^{N}\int d\underline{v}^i_N \gamma_{N-1}(\underline{v}^i_N)\left(\int dv_i e^{-\pi v_i^2} g(v_i) h_N(\underline{v}_N)\right) j_{N-1}(\underline{v}^i_N)\\
&= N((\mathcal{P}^-(g)\mathbf{h})_{N-1}, j_{N-1})_{N-1}\,.
\end{aligned} \tag{52}$$

Assume now that $\mathbf{h}$ is in the domain of $\mathcal{P}^+(g)^*$. This means that for every $\mathbf{j}$ in D^+ we have

$$(\mathcal{P}^+(g)^*\mathbf{h}, \mathbf{j}) = (\mathbf{h}, \mathcal{P}^+(g)\mathbf{j})\,.$$

Given M, choose $\mathbf{j}$ such that $j_N \equiv 0$ if $N \neq M$. For such a $\mathbf{j}$ we have $\mathbf{j} \in D^+$ and

$$\begin{aligned}
\rho a_M((\mathcal{P}^+(g)^*\mathbf{h})_M, j_M)_M &= \rho(\mathcal{P}^+(g)^*\mathbf{h}, \mathbf{j}) = \rho(\mathbf{h}, \mathcal{P}^+(g)\mathbf{j})\\
&= \rho a_{M+1}(h_{M+1}, (\mathcal{P}^+(g)\mathbf{j})_{M+1})_{M+1} = \mu a_M((\mathcal{P}^-(g)\mathbf{h})_M, j_M)_M
\end{aligned}$$

where the last equality follows from (52) and the fact that

$$\rho N a_N = \mu a_{N-1}\,. \tag{53}$$

This implies that $\rho(\mathcal{P}^+(g)^*\mathbf{h})_M = \mu(\mathcal{P}^-(g)\mathbf{h})_M$ for every M thus proving (51). This also implies that $\mathcal{G}$ is self adjoint. ◻

To obtain convergence toward $\mathbf{e}^0$, we first need to show that $\mathcal{G}$ is non positive. This is the content of the following Lemma.

Lemma 11 *$\mathcal{G}$ is non positive and $\mathcal{G}\mathbf{h} = 0$ if and only if $\mathbf{h} = c\mathbf{e}^0$, where $\mathbf{e}^0$ is given by $e^0_N(\underline{v}_N) = 1$ for every N and $\underline{v}_N$.*

Proof From (51), we get $\rho(\mathbf{h}, \mathcal{P}^+\mathbf{h}) = \mu(\mathcal{P}^-\mathbf{h}, \mathbf{h})$ so that

$$(\mathbf{h}, \mathcal{G}\mathbf{h}) = 2\rho(\mathbf{h}, \mathcal{P}^+\mathbf{h}) - (\mathbf{h}, (\rho\mathcal{N} + \mu)\mathbf{h}) \tag{54}$$

Moreover we have

$$\begin{aligned}
\rho(\mathbf{h}, \mathcal{P}^+\mathbf{h}) &= \rho \sum_{N=1}^{\infty} a_N \int d\underline{v}_N \gamma_N(\underline{v}_N) h_N(\underline{v}_N) \left(\sum_{i=1}^{N} h_{N-1}(\underline{v}_N^i) \right) \\
&= \sum_{N=1}^{\infty} \left[\sum_{i=1}^{N} \int d\underline{v}_N \gamma_N(\underline{v}_N) \left(\sqrt{\rho a_N} h_N(\underline{v}_N)\right) \left(\sqrt{\frac{\mu}{N} a_{N-1}} h_{N-1}(\underline{v}_N^i) \right) \right] \\
&\le \sum_{N=1}^{\infty} \sum_{i=1}^{N} \left[\frac{1}{2} \rho a_N \int d\underline{v}_N \gamma_N(\underline{v}_N) h_N(\underline{v}_N)^2 + \frac{1}{2} \frac{\mu}{N} a_{N-1} \int d\underline{v}_N \gamma_N(\underline{v}_N) h_{N-1}(\underline{v}_N^i)^2 \right] \\
&= \sum_{N=0}^{\infty} \left[\frac{1}{2} N \rho a_N \int d\underline{v}_N \gamma_N(\underline{v}_N) h_N(\underline{v}_N)^2 + \frac{1}{2} \mu a_N \int d\underline{v}_N \gamma_N(\underline{v}_N) h_N(\underline{v}_N)^2 \right] \\
&= \frac{1}{2} (\mathbf{h}, (\rho\mathcal{N} + \mu)\mathbf{h})
\end{aligned} \tag{55}$$

where we have used (53) to obtain the second line and that $ab \le (a^2 + b^2)/2$ in going from the second to the third line of (55). Non positivity follows immediately from (54) and (55). Furthermore, we see that the inequality at the end of the second line of (55) becomes an equality if and only if:

$$\sqrt{\rho a_N} h_N(\underline{v}_N) = \sqrt{\frac{\mu}{N} a_{N-1}} h_{N-1}(\underline{v}_N^i)$$

or $h_N(\underline{v}_N) = h_{N-1}(\underline{v}_N^i)$ for every i and N which implies that $h_N \equiv h_0$. □

Our construction of the eigenvalues and eigenvectors of $\mathcal{G}$ is inspired by the construction of the Fock space for a bosonic quantum field theory, see for example Chap. 6 of [17]. The main observation is that the operators $\mathcal{P}^{\pm}(g)$ defined in (50) have the form of the creation and annihilation operators. Since the "ground state" of $\mathcal{G}$ is $\mathbf{e}^0$, as opposed to the state with no particles $\mathbf{n}$, see (64) below, we will introduce the operators $\mathcal{R}^{\pm}(g)$, see (57) below, that can be thought as quasi particle operators, that is operators that create and destroy excitations above the ground state, see for example [1]. The proofs of the Lemmas in the remaining part of this section should be familiar to readers with a background in QFT.

We start with the commutation relations of the operators $\mathcal{P}^{\pm}(g)$ and $\mathcal{N}$. Setting $\{\mathcal{A}, \mathcal{B}\} = \mathcal{A}\mathcal{B} - \mathcal{B}\mathcal{A}$, we obtain the following Lemma.

Lemma 12 *We have*

$$\begin{aligned}
\{\mathcal{P}^+(g_1), \mathcal{P}^-(g_2)\} &= -(g_1, g_2)\mathrm{Id} \\
\{\mathcal{P}^+(g_1), \mathcal{P}^+(g_2)\} &= \{\mathcal{P}^-(g_1), \mathcal{P}^-(g_2)\} = 0 \\
\{\mathcal{N}, \mathcal{P}^{\pm}(g)\} &= \pm\mathcal{P}^{\pm}(g)
\end{aligned}$$

where

$$(g_1, g_2) = \int_{\mathbb{R}} g_1(w) g_2(w) e^{-\pi w^2} dw\,.$$

Proof We first observe that, due to the symmetry of h_N, we have

$$(\mathcal{P}^-(g)\mathbf{h})_N(\underline{v}_N) = \int \gamma(v_{N+1}) g(v_{N+1}) h_{N+1}(\underline{v}_{N+1}) dv_{N+1} := (P^-_N(g) h_{N+1})(\underline{v}_N)$$

while

$$(\mathcal{P}^+(g)\mathbf{h})_N(\underline{v}_N) = \sum_{i=1}^{N} (P^+_{N,i}(g) h_{N-1})(\underline{v}_N)$$

where

$$(P^+_{N,i}(g) h_{N-1})(\underline{v}_N) = h_{N-1}(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_N) g(v_i)\,.$$

Thus we get

$$\begin{aligned}(\mathcal{P}^-(g_1)\mathcal{P}^-(g_2)\mathbf{h})_N(\underline{v}_N) &= (P^-_N(g_1) P^-_{N+1}(g_2) h_{N+2})(\underline{v}_N) \\ &= \int \gamma(v_{N+1})\gamma(v_{N+2}) g_1(v_{N+1}) g_2(v_{N+2}) \\ &\quad \times h_{N+2}(\underline{v}_{N+2}) dv_{N+1} dv_{N+2}\end{aligned}$$

Using again that h_N is symmetric we get $\{\mathcal{P}^-(g_1), \mathcal{P}^-(g_2)\} = 0$. Moreover, we have

$$\begin{aligned}&P^+_{N,i}(g_1) P^+_{N-1,j}(g_2) h_{N-2}(\underline{v}_N) \\ &= \begin{cases} h_{N-2}(v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_{i-1}, v_{i+1}, \ldots, v_N) g_1(v_i) g_2(v_j) & i > j \\ h_{N-2}(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_j, v_{j+2}, \ldots, v_N) g_1(v_i) g_2(v_{j+1}) & i \le j \end{cases}\end{aligned}$$

so that

$$\begin{cases} P^+_{N,i}(g_1) P^+_{N-1,j}(g_2) h_{N-2} = P^+_{N,j}(g_2) P^+_{N-1,i-1}(g_1) h_{N-2} & i > j \\ P^+_{N,i}(g_1) P^+_{N-1,j}(g_2) h_{N-2} = P^+_{N,j+1}(g_2) P^+_{N-1,i}(g_1) h_{N-2} & i \le j\,. \end{cases}$$

Summing over i and j it follows that $\{\mathcal{P}^+(g_1), \mathcal{P}^+(g_2)\} = 0$.

Similarly we have

$$(P^-_N(g_1) P^+_{N+1,N+1}(g_2) h_N)(\underline{v}_N) = h_N(\underline{v}_N) \int g_1(v_{N+1}) g_2(v_{N+1}) \gamma(v_{N+1}) dv_{N+1}$$

while for $i \le N$ we get

$$\begin{aligned}&(P^-_N(g_1) P^+_{N+1,i}(g_2) h_N)(\underline{v}_N) \\ &= g_2(v_i) \int h_N(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_{N+1}) g_1(v_{N+1}) \gamma(v_{N+1}) dv_{N+1} \\ &= (P^+_{N,i}(g_2) P^-_{N-1}(g_1) h_N)(\underline{v}_N)\,.\end{aligned}$$

Summing over i we get $\{\mathcal{P}^+(g_1), \mathcal{P}^-(g_2)\} = -(g_1, g_2)\mathrm{Id}$.

Finally we observe that

$$(\mathcal{P}^-(g)\mathcal{N}\mathbf{h})_N = P^-_N(g)(\mathcal{N}\mathbf{h})_{N+1} = (N+1)(\mathcal{P}^-(g)\mathbf{h})_N = ((\mathcal{N} + \mathrm{Id})\mathcal{P}^+(\mathrm{g})\mathbf{h})_{\mathrm{N}}$$

so that $\{\mathcal{N}, \mathcal{P}^-(g)\} = -\mathcal{P}^-(g)$. The commutation relation for $\mathcal{P}^+$ follows taking the adjoint. ☐

Observe that $\mathcal{P}^-(g)\mathbf{e}^0 = (g,1)\mathbf{e}^0$ while from Lemma 12 it follows that

$$\begin{aligned}\{\mathcal{G}, \mathcal{P}^+(g)\} &= \{\mathcal{P}^+(1), \mathcal{P}^+(g)\} - \rho\{\mathcal{N}, \mathcal{P}^+(g)\} + \mu\{\mathcal{P}^-(1), \mathcal{P}^+(g)\} \\ &= -\rho\mathcal{P}^+(q) + \mu(1,g)\mathrm{Id}\end{aligned} \tag{56}$$

that makes it natural to define the new creation and annihilation operators

$$\begin{aligned}\mathcal{R}^+(g) &= \sqrt{\frac{\rho}{\mu}}\mathcal{P}^+(g) - \sqrt{\frac{\mu}{\rho}}(g,1)\,\mathrm{Id} \\ \mathcal{R}^-(g) &= \sqrt{\frac{\mu}{\rho}}\mathcal{P}^-(g) - \sqrt{\frac{\mu}{\rho}}(g,1)\,\mathrm{Id}\,.\end{aligned} \tag{57}$$

The following Corollary collects the relevant properties of $\mathcal{R}^\pm(g)$.

Corollary 13 *We have* $\mathcal{R}^+(g)^* = \mathcal{R}^-(g)$, $\mathcal{R}^-(g)\mathbf{e}^0 = 0$, *and*

$$\begin{aligned}\{\mathcal{R}^+(g_1), \mathcal{R}^-(g_2)\} &= -(g_1, g_2)\mathrm{Id} \\ \{\mathcal{R}^+(g_1), \mathcal{R}^+(g_2)\} &= \{\mathcal{R}^-(g_1), \mathcal{R}^-(g_2)\} = 0 \\ \{\mathcal{N}, \mathcal{R}^\pm(g)\} &= \pm\left(\mathcal{R}^\pm(g) + \sqrt{\frac{\mu}{\rho}}(g,1)\mathrm{Id}\right)\end{aligned}$$

Moreover we also have

$$\{\mathcal{G}, \mathcal{R}^+(g)\} = -\rho\mathcal{R}^+(g)\,, \qquad \{\mathcal{G}, \mathcal{R}^-(g)\} = \rho\mathcal{R}^-(g)\,. \tag{58}$$

Proof It is easy to verify that $\mathcal{R}^-(g)\mathbf{e}^0 = 0$. Moreover we only need to prove (58) since the other relations are immediate consequences of Lemma 12. From (56) we get

$$\{\mathcal{G}, \mathcal{R}^+(g)\} = \sqrt{\frac{\rho}{\mu}}\{\mathcal{G}, \mathcal{P}^+(g)\} = -\rho\sqrt{\frac{\rho}{\mu}}\mathcal{P}^+(g) + \sqrt{\mu\rho}(g,1)\mathrm{Id} = -\rho\mathcal{R}^+(g)$$

The second equation of (58) follows by taking the adjoint of the first. □

Since K_N preserves the space of polynomials of a given degree, see [4], we choose as an orthonormal basis for $L^2(\mathbb{R}, \gamma(v))$ the polynomials

$$L_n(v) = \frac{1}{\sqrt{n!}}H_n(\sqrt{2\pi}v) \tag{59}$$

where

$$H_n(v) = (-1)^n e^{\frac{v^2}{2}}\frac{d^n}{dv^n}e^{\frac{-v^2}{2}}$$

are the standard Hermite polynomials. For every sequence $\underline{\alpha} = (\alpha_0, \alpha_1, \alpha_2, \ldots)$ such that $\alpha_i \in \mathbb{N}$ and $\lambda(\underline{\alpha}) := \sum_{i=0}^{\infty}\alpha_i < \infty$, we define

$$\mathbf{e}_{\underline{\alpha}} = \prod_{i=0}^{\infty}\frac{(\mathcal{R}_i^+)^{\alpha_i}}{\sqrt{\alpha_i!}}\mathbf{e}^0 \tag{60}$$

where $\mathcal{R}_n^\pm = \mathcal{R}^\pm(L_n)$.

Lemma 14 *The vectors* $\mathbf{e}_{\underline{\alpha}}$ *form an orthonormal basis in* $L^2_s(\mathcal{R}, \boldsymbol{\Gamma})$. *Moreover, we have*

$$\mathcal{G}\mathbf{e}_{\underline{\alpha}} = -\rho\lambda(\underline{\alpha})\mathbf{e}_{\underline{\alpha}}\,. \tag{61}$$

Finally we have $\|\mathcal{K}\mathbf{e}_{\underline{\alpha}}\|_2 < \infty$, *so that* $\mathbf{e}_{\underline{\alpha}} \in D^2$, *for every* $\underline{\alpha}$.

Proof If $n_1 \neq n_2$ and $\alpha_1\alpha_2 \neq 0$, using Corollary 13 we get

$$((\mathcal{R}_{n_1}^+)^{\alpha_1}\mathbf{e}^0, (\mathcal{R}_{n_2}^+)^{\alpha_2}\mathbf{e}^0) = (\mathbf{e}^0, (\mathcal{R}_{n_2}^+)^{\alpha_2}(\mathcal{R}_{n_1}^-)^{\alpha_1}\mathbf{e}^0) = 0$$

while

$$\begin{aligned}((\mathcal{R}_n^+)^{\alpha_1}\mathbf{e}^0, (\mathcal{R}_n^+)^{\alpha_2}\mathbf{e}^0) &= ((\mathcal{R}_n^+)^{\alpha_1-1}\mathbf{e}^0, \mathcal{R}_n^-(\mathcal{R}_n^+)^{\alpha_2}\mathbf{e}^0)\\ &= ((\mathcal{R}_n^+)^{\alpha_1-1}\mathbf{e}^0, \mathcal{R}_n^+\mathcal{R}_n^-(\mathcal{R}_n^+)^{\alpha_2-1}\mathbf{e}^0)\\ &\quad + ((\mathcal{R}_n^+)^{\alpha_1-1}\mathbf{e}^0, (\mathcal{R}_n^+)^{\alpha_2-1}\mathbf{e}^0)\\ &\vdots\\ &= ((\mathcal{R}_n^+)^{\alpha_1-1}\mathbf{e}^0, (\mathcal{R}_n^+)^{\alpha_2}\mathcal{R}_n^-\mathbf{e}^0)\\ &\quad + \alpha_2((\mathcal{R}_n^+)^{\alpha_1-1}\mathbf{e}^0, (\mathcal{R}_n^+)^{\alpha_2-1}\mathbf{e}^0)\\ &= \alpha_2((\mathcal{R}_n^+)^{\alpha_1-1}\mathbf{e}^0, (\mathcal{R}_n^+)^{\alpha_2-1}\mathbf{e}^0)\,.\end{aligned}$$

Assuming $\alpha_1 \geq \alpha_2$ we get

$$((\mathcal{R}_n^+)^{\alpha_1}\mathbf{e}^0, (\mathcal{R}_n^+)^{\alpha_2}\mathbf{e}^0) = \alpha_2!((\mathcal{R}_n^+)^{\alpha_1-\alpha_2}\mathbf{e}^0, \mathbf{e}^0) \tag{62}$$

so that

$$((\mathcal{R}_n^+)^{\alpha_1}\mathbf{e}^0, (\mathcal{R}_n^+)^{\alpha_2}\mathbf{e}^0) = \alpha_1!\delta_{\alpha_1,\alpha_2}$$

from which orthonormality follows easily. Observe now that

$$((\mathcal{P}^+(1))^n\mathbf{e}^0)_N = \begin{cases}0 & N < n\\ \frac{N!}{(N-n)!} & N \geq n\end{cases} \tag{63}$$

so that we can write

$$\mathbf{n} = \sum_{n=0}^{\infty}\frac{(-1)^n}{n!}(\mathcal{P}^+(1))^n\mathbf{e}^0 \tag{64}$$

where $\mathbf{n} = (1, 0, 0, \ldots)$. Since $\mathcal{P}^+(1) = \sqrt{\frac{\rho}{\mu}}\mathcal{R}_0^+ + \sqrt{\frac{\mu}{\rho}}\mathrm{Id}$ we see that $\mathbf{n}$ is in the closure of the span of the $\mathbf{e}_{\underline{\alpha}}$. Calling $\mathcal{P}_i^+ = \mathcal{P}^+(L_i)$, we observe that $(\mathcal{P}_i^+\mathbf{n})_N = 0$ for $N \neq 1$ while $(\mathcal{P}_i^+\mathbf{n})_1 = L_i$. Since the L_i form a basis for $L^2(\mathbb{R}, \gamma_1)$ we see that the closure of the span of $\{\mathbf{n}; \mathcal{P}_i^+\mathbf{n}, i \geq 0\}$ contains a basis for $L_s^2(\mathbb{R}^0, a_0) \oplus L_s^2(\mathbb{R}, a_1\gamma_1)$. Observe now that $\mathcal{P}_i^+ = \sqrt{\frac{\mu}{\rho}}\mathcal{R}_i^+ + \delta_{i,0}\frac{\mu}{\rho}\mathrm{Id}$ and that $\mathcal{R}_i^+\mathbf{e}_{\underline{\alpha}} = \sqrt{\alpha_i + 1}\mathbf{e}_{\underline{\alpha}'}$, where $\alpha_j' = \alpha_j$ for $j \neq i$ while $\alpha_i' = \alpha_i + 1$. Combining this with (64) we get that the closure of the span of the $\mathbf{e}_{\underline{\alpha}}$ contains $\mathcal{P}_i^+\mathbf{n}$ and thus it contains a basis for $L_s^2(\mathbb{R}^0) \oplus L_s^2(\mathbb{R}, a_1\gamma_1)$. Iterating this construction we obtain completeness. Equation (61) follows easily from (58).

Finally, since $(h_N, R_{i,j}h_N)_N \leq \|h_n\|_{2,N}$, from (5) we get

$$\|\mathcal{K}\mathbf{h}\|_2^2 \leq \sum_{N=0}^{\infty} a_N N^4\|\mathbf{h}\|_{2,N}^2 = \|\mathcal{N}^2\mathbf{h}\|_2^2\,.$$

Using the commutation relations in Corollary 13 as in the derivation of (62) we get

$$\mathcal{N}\left(\mathcal{R}_n^+\right)^{\alpha} = \left(\mathcal{R}_n^+\right)^{\alpha}\mathcal{N} + \alpha\left(\mathcal{R}_n^+\right)^{\alpha} + \delta_{n,0}\alpha\sqrt{\frac{\mu}{\rho}}\left(\mathcal{R}_n^+\right)^{\alpha-1}$$

that, together with $\mathcal{N}\mathbf{e}^0 = \sqrt{\frac{\mu}{\rho}}\mathcal{R}_0^+\mathbf{e}^0 + \frac{\mu}{\rho}\mathbf{e}^0$, gives

$$\mathcal{N}\mathbf{e}_{\underline{\alpha}} = \left(\lambda(\underline{\alpha}) + \frac{\mu}{\rho}\right)\mathbf{e}_{\underline{\alpha}} + \sqrt{\frac{\mu}{\rho}}(\sqrt{\alpha_0}\mathbf{e}_{\underline{\alpha}^-} + \sqrt{\alpha_0 + 1}\mathbf{e}_{\underline{\alpha}^+})$$

where $\alpha_i^\pm = \alpha_i$, for $i > 0$, while $\alpha_0^\pm = \alpha_0 \pm 1$. Thus we have $\|\mathcal{N}^2\mathbf{e}_{\underline{\alpha}}\|_2 < \infty$ and the proof is complete. $\square$

In Sect. 3.3 we will need a more explicit representation of the $\mathbf{e}_{\underline{\alpha}}$. To this end observe that, if $n \neq 0$, $(\mathcal{R}_n^+\mathbf{e}^0)_N(\underline{v}_N) = \sqrt{\frac{\rho}{\mu}}\sum_{i=1}^N L_n(v_i)$ while for $n_1, n_2 \neq 0$ and $N \geq 2$ we can write

$$(\mathcal{R}_{n_1}^+\mathcal{R}_{n_2}^+\mathbf{e}^0)_N(\underline{v}_N) = \frac{\rho}{\mu}\sum_{i\neq j} L_{n_1}(v_i)L_{n_2}(v_j) = \frac{1}{(N-2)!}\frac{\rho}{\mu}\sum_{\pi\in\mathrm{Sym}(N)} L_{n_1}(v_{\pi(1)})L_{n_2}(v_{\pi(2)})$$

where $\mathrm{Sym}(N)$ is the group of permutations on $\{1, \ldots, N\}$. More generally, given $n_i \neq 0$, $i = 1, \ldots, M$, we get, for $N \geq M$,

$$\left(\prod_{i=1}^M \mathcal{R}_{n_i}^+\mathbf{e}^0\right)_N(\underline{v}_N) = \frac{1}{(N-M)!}\left(\frac{\rho}{\mu}\right)^{\frac{M}{2}}\sum_{\pi\in\mathrm{Sym}(N)}\prod_{i=1}^M L_{n_i}(v_{\pi(i)})\,. \tag{65}$$

while $\left(\prod_{i=1}^M \mathcal{R}_{n_i}^+\mathbf{e}^0\right)_N \equiv 0$ for $N < M$. Given $\underline{\alpha}$ with $\lambda(\underline{\alpha}) < \infty$, define

$$L_{\underline{\alpha}} = \bigotimes_{i=1}^{\infty} L_i^{\otimes\alpha_i}$$

where $L_i^{\otimes 0} = 1$ and observe that $L_{\underline{\alpha}}$ is a polynomial in $\lambda_0(\underline{\alpha}) := \sum_{i=1}^\infty \alpha_i$ variables with degree $d(\underline{\alpha}) := \sum_{i=1}^\infty i\alpha_i$. Also for $\pi \in \mathrm{Sym}(N)$, define $\pi(\underline{v}_N) = (v_{\pi(1)}, v_{\pi(2)}, \ldots v_{\pi(N)})$. Using these definitions, together with (60) and the fact that $\mathcal{R}_0^+ = \sqrt{\frac{\rho}{\mu}}\mathcal{P}^+(1) + \sqrt{\frac{\mu}{\rho}}\mathrm{Id}$ we can write, for $N \geq \lambda_0(\underline{\alpha})$,

$$(\mathbf{e}_{\underline{\alpha}})_N(\underline{v}_N) = c_{\underline{\alpha},N}\sum_{\pi\in\mathrm{Sym}(N)} L_{\underline{\alpha}}(\pi(\underline{v}_N))\,, \tag{66}$$

for suitable coefficients $c_{\underline{\alpha},N}$, while $(\mathbf{e}_{\underline{\alpha}})_N(\underline{v}_N) = 0$ for $N < \lambda_0(\underline{\alpha})$.

We now come back to the full operator $\widetilde{\mathcal{L}}$.

Corollary 15 *The operator $\widetilde{\mathcal{L}}$ is self-adjoint, non positive and $\widetilde{\mathcal{L}}\mathbf{h} = 0$ if and only if $\mathbf{h} = c\mathbf{e}^0$.*

Proof We can proceed exactly as in proof of Lemma 10. Assume that $\mathbf{h}$ is in the domain of $\widetilde{\mathcal{L}}^*$. This means that for every $\mathbf{j}$ in D^2 we have

$$(\widetilde{\mathcal{L}}^*\mathbf{h}, \mathbf{j}) = (\mathbf{h}, \widetilde{\mathcal{L}}\mathbf{j})\,.$$

Given M, choose $\mathbf{j}$ such that $j_N \equiv 0$ if $N \neq M$. Clearly $\mathbf{j} \in D^2$ because $(\widetilde{\mathcal{L}}\mathbf{j})_N \neq 0$ only for $N = M-1$, M, and $M+1$. Moreover $(\mathcal{K}\mathbf{h}, \mathbf{j}) = a_M(K_M h_M, j_M)_M$ is well defined for every $\mathbf{h} \in L_s^2(\mathcal{R}, \boldsymbol{\Gamma})$. Finally we known that K_M is non negative and self-adjoint for every M. Thus we get

$$\begin{aligned} a_M((\widetilde{\mathcal{L}}^*\mathbf{h})_M, j_M)_M &= ((\widetilde{\mathcal{L}}^*\mathbf{h}, \mathbf{j}) = (\mathbf{h}, \widetilde{\mathcal{L}}\mathbf{j}) = (\mathbf{h}, \mathcal{G}\mathbf{j}) + \tilde{\lambda}a_M(h_M, K_M j_M)_M \\ &= a_M((\mathcal{G}\mathbf{h})_M, j_M) + \tilde{\lambda}a_M(K_M h_M, j_M)_M = a_M((\widetilde{\mathcal{L}}\mathbf{h})_M, j_M)_M\,. \end{aligned}$$

This implies that $(\widetilde{\mathcal{L}}^*\mathbf{h})_M = (\widetilde{\mathcal{L}}\mathbf{h})_M$ for every M. This proves that $\widetilde{\mathcal{L}}$ is self-adjoint. Observe also that $\mathcal{G}\mathbf{h} = 0$ if and only if $\mathbf{h} = c\mathbf{e}^0$, see Lemma 11, while $\mathcal{K}$ is positive and $\mathcal{K}\mathbf{e}^0 = 0$. This completes the proof. □

Let $\mathbf{W}_1 = \text{span}\{\mathbf{e}_{\underline{\alpha}} \,|\, \lambda(\underline{\alpha}) = 1\} = \text{span}\{\mathcal{R}_n^+\mathbf{e}^0 \,|\, n \geq 0\}$. Observe that $\mathcal{G}\mathbf{h} = -\rho\mathbf{h}$ if $\mathbf{h} \in \mathbf{W}_1$ while $(\mathbf{h}, \mathcal{G}\mathbf{h}) < -\rho(\mathbf{h}, \mathbf{h})$ if $\mathbf{h} \in D^2$, $\mathbf{h} \perp \mathbf{e}^0$ but $\mathbf{h} \notin \mathbf{W}_1$. Thus we get

$$\Delta \leq -\rho + \sup\{(\mathbf{h}, \mathcal{K}\mathbf{h}) \,|\, \mathbf{h} \in D^2, \|\mathbf{h}\|_2 = 1, \mathbf{h} \perp \mathbf{E_0}\} \leq -æ$$

From [4] we know that $(f_N, K_N f_N) \leq 0$ for every f_N while $(f_N, K_N f_N) = 0$ if and only if f_N is rotationally invariant. Since $(\mathcal{R}_n^+\mathbf{e}^0)_N = \sqrt{\rho/\mu}\sum_{i=1}^N L_n(v_i)$, for $n > 0$, while $(\mathcal{R}_0^+\mathbf{e}^0)_N = \sqrt{\rho/\mu}N - \sqrt{\mu/\rho}$ we have that $\mathcal{R}_n^+\mathbf{e}^0$ is rotationally invariant if and only if $n = 0$ or $n = 2$. This implies that $(\mathbf{h}, \widetilde{\mathcal{L}}\mathbf{h}) = -\rho\|\mathbf{h}\|_2$ if and only if $\mathbf{h} \in \text{span}\{\mathcal{R}_0^+\mathbf{e}^0, \mathcal{R}_2^+\mathbf{e}^0\}$. Since $\mathcal{R}_0^+\mathbf{e}^0 = \mathbf{e}_{(1,0,\ldots)}$ and $\mathcal{R}_2^+\mathbf{e}^0 = \mathbf{e}_{(0,0,1,0,\ldots)}$, this completes the proof of Theorem 2. □

3.3 Proof of Theorem 3

To prove Theorem 3, we need more information on the action of $\mathcal{K}$ on the basis vectors $\mathbf{e}_{\underline{\alpha}}$.

As a basic step, we compute the action of $R_{1,2}$, see (6), on the product of two Hermite polynomials in v_1 and v_2. A simple calculation, see e.g. [4], shows that $(R_{1,2}F)(v_1, v_2) = 0$ for every F odd in v_1 or v_2. Thus, calling $H_{(m_1,m_2)}(v_1, v_2) = H_{m_1}(v_1)H_{m_2}(v_2)$, it follows that $R_{1,2}H_{(m_1,m_2)} \neq 0$ if and only if m_1 and m_2 are both even while $R_{1,2}H_{(2n_1,2n_2)}$ is a rotationally invariant polynomial of degree $2(n_1 + n_2)$ in v_1 and v_2. Moreover, if $m_1 + m_2 < 2n_1 + 2n_2$, we get

$$\begin{aligned}
&\int H_{(m_1,m_2)}(v_1, v_2)\big(R_{1,2}H_{(2n_1,2n_2)}\big)(v_1, v_2)\gamma(v_1)\gamma(v_2)dv_1dv_2 \\
&= \int \big(R_{1,2}H_{(m_1,m_2)}\big)(v_1, v_2)H_{(2n_1,2n_2)}(v_1, v_2)\gamma(v_1)\gamma(v_2)dv_1dv_2 = 0
\end{aligned}$$

where we have used that $H_{(2n_1,2n_2)}$ is orthogonal to any polynomial of degree less that $2(n_1 + n_2)$. Thus we have $R_{1,2}H_{(2n_1,2n_2)} \in \text{span}\{H_{(p_1,p_2)} \,|\, p_1 + p_2 = 2n_1 + 2n_2\}$ and, since H_n is a monic polynomial of degree n, we can write

$$R_{1,2}H_{(2n_1,2n_2)} = \sum_{k=0}^{n_1+n_2} a_{k,n_1,n_2}H_{(2k,2(n_1+n_2-k))} = \sum_{k=0}^{n_1+n_2} a_{k,n_1,n_2}v_1^{2k}v_2^{2(n_1+n_2-k)} + Q$$

for suitable coefficients a_{k,n_1,n_2} and polynomial $Q(v_1, v_2)$ of degree strictly less then $2(n_1 + n_2)$. This, together with rotational invariance, implies that

$$R_{1,2}H_{(2n_1,2n_2)} = \tilde{\tau}_{n_1,n_2}\sum_{k=0}^{n_1+n_2}\binom{n_1+n_2}{k}H_{(2k,2(n_1+n_2-k))} \tag{67}$$

for suitable coefficients $\tilde{\tau}_{n,m}$. Using (67), together with (66), it is possible to give an explicit representation of $\mathcal{K}$ on the basis of the $\mathbf{e}_{\underline{\alpha}}$. For the purpose of this paper, we will only need some particular cases discussed in detail below.

Let now $\mathbf{V}_m = \text{span}\{\mathbf{e}_{\underline{\alpha}} | \sum_{i=1}^\infty i\alpha_i = m\} = \text{span}\{\prod_i(\mathcal{R}_i^+)^{\alpha_i}\mathbf{e}^0 | \sum_{i=1}^\infty i\alpha_i = m\}$, that is $\mathbf{V}_m$ is the subspace of all states $\mathbf{h}$ such that h_N is a polynomial of degree m orthogonal to all polynomials of degree less than m. From the above considerations and (66) it follows that

$\mathcal{K}\mathbf{V}_m \subset \mathbf{V}_m$ so that defining

$$\delta_m = \inf_{\substack{\mathbf{h}\in\mathbf{V}_m\cap D^2 \\ \|\mathbf{h}\|_2=1,\, \mathbf{h}\perp\mathbf{E}_1\oplus\mathbf{E}_0}} (\mathbf{h}, -\widetilde{\mathcal{L}}\mathbf{h})\,. \tag{68}$$

and observing that $L^2_s(\mathcal{R}, \boldsymbol{\Gamma}) = \bigoplus_{m=0}^{\infty} \mathbf{V}_m$, we get $\Delta_2 = -\inf_m \delta_m$.

Since $\mathbf{E}_1 = \mathrm{span}\{\mathcal{R}_0^+\mathbf{e}^0, \mathcal{R}_2^+\mathbf{e}^0\}$, we get

$$\begin{aligned}
\mathbf{V}_0 \cap (\mathbf{E}_1 \oplus \mathbf{E}_0)^\perp &= \mathrm{span}\{(\mathcal{R}_0^+)^n\mathbf{e}^0, n \geq 2\} \\
\mathbf{V}_2 \cap (\mathbf{E}_1 \oplus \mathbf{E}_0)^\perp &= \mathrm{span}\{(\mathcal{R}_0^+)^n\mathcal{R}_2^+\mathbf{e}^0, n \geq 1;\, (\mathcal{R}_0^+)^m(\mathcal{R}_1^+)^2\mathbf{e}^0, m \geq 0\}\,.
\end{aligned}$$

Observing that $\mathcal{K}(\mathcal{R}_0^+)^n\mathbf{e}^0 = \mathcal{K}(\mathcal{R}_0^+)^n\mathcal{R}_2^+\mathbf{e}^0 = 0$, due to rotational invariance, while $\mathcal{K}(\mathcal{R}_0^+)^m(\mathcal{R}_1^+)^2\mathbf{e}^0 = 0$, due to parity, we obtain $\delta_0 = \delta_2 = 2\rho$. Moreover we have that, for $m \neq 0, 2$, $\mathbf{V}_m \perp \mathbf{E}_1 \oplus \mathbf{E}_0$. Thus we need a lower bound on δ_m for m odd and for m even and greater than 2.

Observe that $(\mathcal{R}_m^+\mathbf{e}^0, \mathcal{G}\mathcal{R}_m^+\mathbf{e}^0) = -\rho$ while $(\mathbf{h}, \mathcal{G}\mathbf{h}) \leq -2\rho(\mathbf{h}, \mathbf{h})$ if $\mathbf{h} \in \mathbf{V}_m$ and $\mathbf{h} \perp \mathcal{R}_m^+\mathbf{e}^0$. Thus, if λ is not too big, it is natural to search for the infimum of $(\mathbf{h}, -\widetilde{\mathcal{L}}\mathbf{h})$ on $\mathbf{V}_m$ looking at states $\mathbf{h}$ close to $\mathcal{R}_m^+\mathbf{e}^0$. To do this, we need the representation of $\mathcal{K}\mathcal{R}_m^+\mathbf{e}^0$ on the basis formed by the $\mathbf{e}_{\underline{\alpha}}$. If $m = 2n$, using (67) for $n_2 = 0$ we get

$$R_{1,2}H_{(2n,0)} = \tau_n \sum_{k=0}^{n} \binom{n}{k} H_{(2k,2(n-k))} \tag{69}$$

where $\tau_n = \tilde{\tau}_{n,0}$. To compute τ_n we compare the coefficients of v_1^{2n} on the left and right hand side of (69). On the left hand side the only contribution comes from $R_{1,2}v_1^{2n}$ since $R_{1,2}$ preserve the degree. On the right hand side only the term with $k = n$ contains the monomial v_1^{2n}. Since the H_n are monic and

$$R_{1,2}v_1^{2n} = \int_0^{2\pi} (v_1\cos\theta - v_2\sin\theta)^{2n}\frac{d\theta}{2\pi} = (v_1^2 + v_2^2)^n \int_0^{2\pi} \cos^{2n}\theta\,\frac{d\theta}{2\pi}\,,$$

and we obtain

$$\tau_n = \int_0^{2\pi} \cos^{2n}\theta\,\frac{d\theta}{2\pi} = \frac{1}{4^n}\binom{2n}{n}.$$

Combining with (59) we get

$$R_{i,j}L_{2n}(v_i) = \tau_n \sum_{k=0}^{n} \binom{n}{k} \frac{\sqrt{(2k)![2(n-k)]!}}{\sqrt{(2n)!}} L_{2k}(v_i)L_{2(n-k)}(v_j)\,.$$

Since for $n > 0$ we have $(\mathcal{R}_{2n}^+\mathbf{e}^0)_N = \sqrt{\rho/\mu}\sum_{i=1}^N L_{2n}(v_i)$, a direct computation shows that

$$\begin{aligned}
(\mathcal{K}\mathcal{R}_{2n}^+\mathbf{e}^0)_N &= \sqrt{\frac{\rho}{\mu}}(N-1)(2\tau_n - 1)\sum_{i=1}^{N} L_{2n}(v_i) \\
&\quad + \sqrt{\frac{\rho}{\mu}}\sum_{k=1}^{n-1}\sum_{i\neq j}\sigma_{n,k}L_{2k}(v_i)L_{2(n-k)}(v_j)
\end{aligned}$$

where

$$\sigma_{n,k} = \tau_n \frac{\binom{n}{k}}{\sqrt{\binom{2n}{2k}}} = \sqrt{\tau_n\tau_k\tau_{n-k}}\,. \tag{70}$$

This gives us

$$
\begin{aligned}
\mathcal{K}\mathcal{R}_{2n}^{+}\mathbf{e}^{0} &= (2\tau_n - 1)\mathcal{R}_{2n}^{+}\mathcal{N}\mathbf{e}^{0} + \sqrt{\frac{\mu}{\rho}}\sum_{k=1}^{n-1}\sigma_{n,k}\mathcal{R}_{2k}^{+}\mathcal{R}_{2(n-k)}^{+}\mathbf{e}^{0} \\
&= \frac{\mu}{\rho}(2\tau_n - 1)\mathcal{R}_{2n}^{+}\mathbf{e}^{0} + \sqrt{\frac{\mu}{\rho}}(2\tau_n - 1)\mathcal{R}_{0}^{+}\mathcal{R}_{2n}^{+}\mathbf{e}^{0} \\
&\quad + \sqrt{\frac{\mu}{\rho}}\sum_{k=1}^{n-1}\sigma_{n,k}\mathcal{R}_{2k}^{+}\mathcal{R}_{2(n-k)}^{+}\mathbf{e}^{0}
\end{aligned}
\tag{71}
$$

where we have used that $\mathcal{N}\mathbf{e}^{0} = \sqrt{\frac{\mu}{\rho}}\mathcal{R}_{0}^{+}\mathbf{e}^{0} + \frac{\mu}{\rho}\mathbf{e}^{0}$.

If $m = 2n + 1$, $R_{1,2}H_{2n+1}(v_1) = 0$ gives

$$
\mathcal{K}\mathcal{R}_{2n+1}^{+}\mathbf{e}^{0} = -\frac{\mu}{\rho}\mathcal{R}_{2n+1}^{+}\mathbf{e}^{0} - \sqrt{\frac{\mu}{\rho}}\mathcal{R}_{0}^{+}\mathcal{R}_{2n+1}^{+}\mathbf{e}^{0}\,. \tag{72}
$$

From (71) and (72) we get

$$
\tilde{\lambda}(\mathcal{R}_{2n}^{+}\mathbf{e}^{0}, \mathcal{K}\mathcal{R}_{2n}^{+}\mathbf{e}^{0}) = -\lambda(1 - 2\tau_n)\,, \quad \tilde{\lambda}(\mathcal{R}_{2n+1}^{+}\mathbf{e}^{0}, \mathcal{K}\mathcal{R}_{2n+1}^{+}\mathbf{e}^{0}) = -\lambda
$$

so that $\delta_{2n} \le \rho + \lambda(1 - 2\tau_n)$ and $\delta_{2n+1} \le \rho + \lambda$.

The following Lemma shows that, if the average number of particles in the steady state is large enough and λ is not too large, one can find a lower bound for δ_m close to the upper bound derived above.

Lemma 16 *For $m = 2n + 1$ we have*

$$
\delta_{2n+1} \ge \min\left\{\rho + \lambda - \lambda\sqrt{\frac{\rho}{\mu}},\ 2\rho - \lambda\sqrt{\frac{\rho}{\mu}}\right\} \tag{73}
$$

while for $m = 2n$, $n > 1$, we have

$$
\delta_{2n} \ge \min\left\{\rho + (1 - 2\tau_n)\lambda - 2\lambda\sqrt{\frac{\rho}{\mu}},\ 2\rho - 2\lambda\sqrt{\frac{\rho}{\mu}}\right\}\,. \tag{74}
$$

Proof See Appendix A.1. □

Since $\tau_2 = 3/8$ and $(\mathcal{R}_4\mathbf{e}^{0}, -\widetilde{\mathcal{L}}\mathcal{R}_4\mathbf{e}^{0}) = \rho + \lambda/4$, we get

$$
\rho + \frac{\lambda}{4} - 2\lambda\sqrt{\frac{\rho}{\mu}} < \delta_4 \le \rho + \lambda/4.
$$

Moreover, thanks to (13),

$$
2\rho - \lambda\sqrt{\frac{\rho}{\mu}} > \rho + \frac{\lambda}{4}\,, \qquad \rho + \lambda - \lambda\sqrt{\frac{\rho}{\mu}} > \rho + \frac{\lambda}{4}
$$

so that $\delta_{2n+1} > \delta_4$ for every n. Finally we observe that $\tau_{n+1} < \tau_n$ and $\tau_3 = 5/16$. Using (13) again it follows that, for $n \ge 3$,

$$
\delta_{2n} \ge \min\left\{2\rho - 2\lambda\sqrt{\frac{\rho}{\mu}}, (1 - 2\tau_3)\lambda + \rho - 2\lambda\sqrt{\frac{\rho}{\mu}}\right\} > \rho + \frac{\lambda}{4} \ge \delta_4
$$

so that $\Delta_2 = -\delta_4$.

To show that Δ_2 is an eigenvalue, we need to construct an eigenstate, that is we need to find $\hat{\mathbf{h}} \in \mathbf{V}_4$ such that $\widetilde{\mathcal{L}}\hat{\mathbf{h}} = -\delta_4\hat{\mathbf{h}}$. To this end, it is enough to show that there exists $\hat{\mathbf{h}} \in \mathbf{V}_4$

such that $(\hat{\mathbf{h}}, \widetilde{\mathcal{L}}\hat{\mathbf{h}}) = -\delta_4(\hat{\mathbf{h}}, \hat{\mathbf{h}})$. Observe that if $\mathbf{h} \in \mathbf{V}_4$ then $\mathcal{K}\mathbf{h}$ is even. We thus restrict our search to $\hat{\mathbf{h}} \in \mathbf{V}_4^e = \mathrm{span}\{(\mathcal{R}_0^+)^k\mathcal{R}_4^+\mathbf{e}^0,\ (\mathcal{R}_0^+)^k(\mathcal{R}_2^+)^2\mathbf{e}^0;\ k \geq 0\}$.

Consider a sequence $\mathbf{h}_n \in \mathbf{V}_4^e$ such that $\|\mathbf{h}_n\|_2 = 1$ and $\lim_{n\to\infty}(\mathbf{h}_n, -\widetilde{\mathcal{L}}\mathbf{h}_n) = \delta_4$. Calling $\mathbf{V}_{4,k}^e = \mathrm{span}\{(\mathcal{R}_0^+)^{k-1}\mathcal{R}_4^+\mathbf{e}^0, (\mathcal{R}_0^+)^{k-2}(\mathcal{R}_2^+)^2\mathbf{e}^0\}$ for $k > 2$, while $\mathbf{V}_{4,1}^e = \mathrm{span}\{\mathcal{R}_4^+\mathbf{e}^0\}$, we can write $\mathbf{h}_n = \sum_{k=0}^{\infty}\mathbf{h}_{n,k}$ with $\mathbf{h}_{n,k} \in \mathbf{V}_{4,k}^e$ and we can find a subsequence $\mathbf{h}_n^0$ of $\mathbf{h}_n$ such that $\lim_{n\to\infty}\mathbf{h}_{n,0} = \hat{\mathbf{h}}_0$. Similarly we can find a new subsequence $\mathbf{h}_n^1$ of $\mathbf{h}_n^0$ such that $\lim_{n\to\infty}\mathbf{h}_{n,1} = \hat{\mathbf{h}}_1$. Proceeding like this we find a sequence $\mathbf{h}_n^\infty$ such that $\lim_{n\to\infty}\mathbf{h}_{n,k}^\infty = \hat{\mathbf{h}}_k$, for every k. Analogously, since $h_{n,N}$ is an even polynomial of degree 4 in $\underline{v}_N$ we can assume, possibly at the cost of further extracting a subsequence, that $\lim_{n\to\infty} h_{n,N}^\infty = \hat{h}_N$ for every N. From Fatou's Lemma we get that $\lim_{n\to\infty}\mathbf{h}_n^\infty = \hat{\mathbf{h}}$ with $\|\hat{\mathbf{h}}\|_2 \leq 1$ while

$$(\hat{\mathbf{h}}, -\mathcal{G}\hat{\mathbf{h}}) = \rho\sum_{k=1}^{\infty}k\|\hat{\mathbf{h}}_k\|^2 \leq \liminf_{n\to\infty}\rho\sum_{k=1}^{\infty}k\|\mathbf{h}_{n,k}\|^2 = \liminf_{n\to\infty}(\mathbf{h}_n^\infty, -\mathcal{G}\mathbf{h}_n^\infty)$$

and analogously, since K_N is non positive,

$$\begin{aligned}(\hat{\mathbf{h}}, -\mathcal{K}\hat{\mathbf{h}}) &= \sum_{N=0}^{\infty}(\hat{h}_N, -K_N\hat{h}_N)_N \leq \liminf_{n\to\infty}\sum_{N=0}^{\infty}(h_{n,N}^\infty, -K_N h_{n,N}^\infty)_N \\ &\leq \liminf_{n\to\infty}(\mathbf{h}_n^\infty, -\mathcal{K}\mathbf{h}_n^\infty)\end{aligned}$$

so that

$$(\hat{\mathbf{h}}, -\widetilde{\mathcal{L}}\hat{\mathbf{h}}) \leq \liminf_{n\to\infty}(\mathbf{h}_n^\infty, -\widetilde{\mathcal{L}}\mathbf{h}_n^\infty) = \delta_4$$

while $(\hat{\mathbf{h}}, -\widetilde{\mathcal{L}}\hat{\mathbf{h}}) \geq \delta_4\|\hat{\mathbf{h}}\|_2$ since $\hat{\mathbf{h}} \in \mathbf{V}_4^e$. Thus we need to show that $\|\hat{\mathbf{h}}\|_2 = 1$.

To this end observe that for every $M > 0$ we have

$$\rho M\sum_{k=M+1}^{\infty}\|\mathbf{h}_{n,k}\|_2^2 \leq \rho\sum_{k=1}^{\infty}k\|\mathbf{h}_{n,k}\|_2^2 = (\mathbf{h}_n, -\mathcal{G}\mathbf{h}_n) \leq (\mathbf{h}_n, -\widetilde{\mathcal{L}}\mathbf{h}_n) \leq 2\delta_4$$

eventually in n. Thus, for every ϵ there exists M such that $\sum_{k=0}^{M}\|\mathbf{h}_{n,k}\|_2^2 \geq 1-\epsilon$ eventually in n. Taking the limit this implies that for every ϵ there exists M such that $\sum_{k=0}^{M}\|\hat{\mathbf{h}}_k\|_2^2 \geq 1-\epsilon$ and thus we get $\|\hat{\mathbf{h}}\| = 1$. This concludes the proof of Theorem 3. □

3.4 Proof of Theorem 4

To simplify notation, given $\mathbf{f} = \mathbf{h}\boldsymbol{\Gamma}$, we set $S(\mathbf{h}) = \mathcal{S}(\mathbf{f} \mid \boldsymbol{\Gamma})$ and we define

$$\begin{aligned}\Psi(\mathbf{h}) &= \sum_{N=0}^{\infty}a_N\int d\underline{v}_{N+1}(h_{N+1} - h_N)(\log h_{N+1} - \log h_N)\gamma_{N+1}(\underline{v}_{N+1}) \\ E(\mathbf{h}) &= \sum_{N=0}^{\infty}a_N\int d\underline{v}_N h_N(\underline{v}_N)\gamma_N(\underline{v}_N).\end{aligned}$$

Finally we observe that if $\mathbf{f} \in L_s^1(\mathcal{R})$ then $\mathbf{h} \in L_s^1(\mathcal{R}, \boldsymbol{\Gamma})$ and $e^{\mathcal{L}t}\mathbf{f} = (e^{\widetilde{\mathcal{L}}t}\mathbf{h})\boldsymbol{\Gamma}$ with $\widetilde{\mathcal{L}} = \mathcal{G} + \tilde{\lambda}\mathcal{K}$ defined in Sect. 3.2 but now considered as an operator on $L_s^1(\mathcal{R}, \boldsymbol{\Gamma})$.

To obtain an explicit expression for $\frac{d}{dt}S(\mathbf{h}(t))$, where $\mathbf{h}(t) = e^{\widetilde{\mathcal{L}}t}\mathbf{h}$ we need to interchange the order of the derivative in t with the sum over N and the integral over $\underline{v}_N$. To do this we will use the following two Lemmas that will allow us to use Fatou's Lemma to excahnge derivative and integrals.

Lemma 17 *Given* $\mathbf{f} \in L^1(\mathcal{R})$ *we have*

$$\lim_{t\to 0^+} \left(\left(e^{\mathcal{L}t}\mathbf{f}\right)_N (\underline{v}_N) - f_N(\underline{v}_N)\right) = 0$$
$$\lim_{t\to 0^+} \frac{1}{t} \left(\left(e^{\mathcal{L}t}\mathbf{f}\right)_N (\underline{v}_N) - f_N(\underline{v}_N)\right) = (\mathcal{L}\mathbf{f})_N\,(\underline{v}_N)$$

for every N *and almost every* $\underline{v}_N$.

Proof See Appendix A.2. □

Lemma 18 *If* $\mathbf{h}\boldsymbol{\Gamma} \in L^1_s(\mathcal{R})$ *then*

$$h_N(t)\log(h_N(t)) \le \left(e^{\widetilde{\mathcal{L}}t}(\mathbf{h}\log\mathbf{h})\right)_N .$$

Proof See Appendix A.3. □

After setting

$$\frac{d_+}{dt}S(\mathbf{h}(t)) := \limsup_{h\to 0^+} \frac{1}{h}(S(\mathbf{h}(t+h)) - S(\mathbf{h}(t)))\,,$$

we are ready to estimate of the variation in time of $S(\mathbf{h})$.

Lemma 19 *Let* $\mathbf{h}$ *be such that* $\mathbf{h}\boldsymbol{\Gamma} \in D^1$ *and* $\mathbf{h}\log\mathbf{h}\boldsymbol{\Gamma} \in D^1$ *then we have*

$$\frac{d_+}{dt}S(\mathbf{h}(t)) \le -\mu\Psi(\mathbf{h}(t))\,.$$

Proof From Lemma 18 we get

$$\frac{1}{t}\left(h_N(\underline{v}_N,t)\log(h_N(\underline{v}_N,t)) - h_N(\underline{v}_N)\log(h_N(\underline{v}_N))\right)$$
$$-\frac{1}{t}\left(\left(e^{\widetilde{\mathcal{L}}t}(\mathbf{h}\log\mathbf{h})\right)_N(\underline{v}_N) - h_N(\underline{v}_N)\log(h_N(\underline{v}_N))\right) \le 0\,.$$

Since $\mathbf{h}\log\mathbf{h}\boldsymbol{\Gamma} \in L^1(\mathcal{R})$, conservation of probability gives

$$\sum_{N=0}^{\infty} a_N \int_{\mathbb{R}^N} \left(\left(e^{\widetilde{\mathcal{L}}t}(\mathbf{h}\log\mathbf{h})\right)_N(\underline{v}_N)\gamma_N(\underline{v}_N) - h_N(\underline{v}_N)\log(h_N(\underline{v}_N))\gamma_N(\underline{v}_N)\right) d\underline{v}_N = 0$$

so that by Fatou's Lemma

$$\begin{aligned}
&\limsup_{t\to 0^+}\frac{1}{t}(S(\mathbf{h}(t)) - S(\mathbf{h}))\\
&\le \sum_{N=0}^{\infty} a_N \int_{\mathbb{R}^N} \limsup_{t\to 0^+}\frac{1}{t}\left(h_N(\underline{v}_N,t)\log(h_N(\underline{v}_N,t)) - h_N(\underline{v}_N)\log(h_N(\underline{v}_N))\right)\\
&\quad - \sum_{N=0}^{\infty} a_N \int_{\mathbb{R}^N} \limsup_{t\to 0^+}\frac{1}{t}\left(\left(e^{\widetilde{\mathcal{L}}t}(\mathbf{h}\log\mathbf{h})\right)_N(\underline{v}_N) - h_N(\underline{v}_N)\log(h_N(\underline{v}_N))\right)
\end{aligned}$$

and, using Lemma 17, we get

$$\limsup_{t\to 0^+}\frac{1}{t}(S(\mathbf{h}(t)) - S(\mathbf{h})) \le \sum_{N=0}^{\infty} a_N \int_{\mathbb{R}^N} (\widetilde{\mathcal{L}}\mathbf{h})_N(\underline{v}_N)(\log(h_N(\underline{v}_N)) + 1)\gamma_N(\underline{v}_N)d\underline{v}_N$$

$$-\sum_{N=0}^{\infty} a_N \int_{\mathbb{R}^N} \left(\widetilde{\mathcal{L}}(\mathbf{h}\log\mathbf{h})\right)_N (\underline{v}_N)\gamma_N(\underline{v}_N)d\underline{v}_N \,.$$

Since $\boldsymbol{\Gamma}\mathbf{h} \in D^1$ and $\boldsymbol{\Gamma}\mathbf{h}\log\mathbf{h} \in D^1$, (45) gives

$$\begin{aligned}
\frac{d_+}{dt} S(\mathbf{h}(t))\big|_{t=0} &\le \sum_{N=0}^{\infty} a_N \int d\underline{v}_N \gamma_N (\widetilde{\mathcal{L}}\mathbf{h})_N \log(h_N) \\
&= \sum_{N=0}^{\infty} a_N \int d\underline{v}_N \gamma_N \Big(\rho(\mathcal{P}^+\mathbf{h})_N + \mu(\mathcal{P}^-\mathbf{h})_N \\
&\quad -(\mu+\rho N)h_N + \tilde{\lambda} K_N h_N\Big) \log h_N \\
&\le \sum_{N=0}^{\infty} a_N \int d\underline{v}_N \gamma_N \left(\rho(\mathcal{P}^+\mathbf{h})_N + \mu(\mathcal{P}^-\mathbf{h})_N - (\mu+\rho N)h_N\right) \log h_N
\end{aligned}$$

where we have used that $\int d\underline{v}_N \gamma_N (K_N h_N) \log h_N \le 0$. Observe finally that

$$\int d\underline{v}_N \gamma_N \rho (\mathcal{P}^+\mathbf{h})_N \log h_N = \int d\underline{v}_N \gamma_N N \rho h_{N-1} \log h_N$$
$$\int d\underline{v}_N \gamma_N \mu (\mathcal{P}^-\mathbf{h})_N \log h_N = \int d\underline{v}_{N+1} \gamma_{N+1} \mu h_{N+1} \log h_N$$

from which we get

$$\begin{aligned}
\frac{d_+}{dt} S(\mathbf{h}(t))\big|_{t=0} &\le \sum_{N=1}^{\infty} a_N \int d\underline{v}_N \gamma_N N\rho(\tilde{h}_{N-1}(t) - \tilde{h}_N(t)) \log \tilde{h}_N(t) \\
&\quad + \mu \sum_{N=0}^{\infty} a_N \int d\underline{v}_{N+1} \gamma_{N+1} (\tilde{h}_{N+1}(t) - \tilde{h}_N(t)) \log \tilde{h}_N(t)
\end{aligned}$$

The thesis follows by reindexing the first sum and using (53). □

Thus to show that $S(\mathbf{h}(t))$ decays exponentially we need a lower bound for $\Psi(\mathbf{h})$ in terms of $S(\mathbf{h})$. This is the content of the following Lemma that is the main result of this section.

Lemma 20 *If $\boldsymbol{\Gamma}\mathbf{h} \in L^1_s(\mathcal{R})$ with $S(\mathbf{h}) < \infty$, then*

$$S(\mathbf{h}) \le E(\mathbf{h}) \log E(\mathbf{h}) + \frac{\mu}{\rho} \Psi(\mathbf{h}) \,. \tag{75}$$

Remark 21 The idea behind the proof of (75) is to think of the entry and exit processes defined by the thermostat as a continuous family of independent entry processes, one for each possible velocity v, with entry rates $\mu\gamma(v)dv$, while each particle in the system leaves with rate ρ independent of its velocity. Clearly such a description makes little mathematical sense and, as a first step, one may think of approximating the original process by restricting the velocity of each particle to assume only a finite number of values $\bar{v}_k$, $k = 1, \ldots, K$, characterized by suitable entry rates ω_k. After this, using convexity, we reduce the proof of (75) to the case with $K = 1$, essentially equivalent to the case in which all particles in the thermostat have the same velocity. In this situation, we further approximate the infinite reservoir by a large finite reservoir containing M particles that enter and leave the system, independently from each other, at a suitable rate. Convexity will allow us to reduce this situation to that of a single particle jumping from the system to the reservoir and back. The final step is thus

Lemma 25 below that deals with this situation. This argument is inspired by the proof of the Logarithmic Sobolev Inequality in [12].

Remark 22 In the proof of Lemma 19 we required that $\mathbf{h}\boldsymbol{\Gamma} \in D^1$ and $\mathbf{h}\log\mathbf{h}\boldsymbol{\Gamma} \in D^1$ only to differentiate $e^{t\widetilde{\mathcal{L}}}\mathbf{h}$ and show that $\sum_{N=0}^{\infty} a_N \int (\widetilde{\mathcal{L}}\mathbf{h})_N \gamma_N d\underline{v}_N = 0$ and similarly for $\mathbf{h}\log\mathbf{h}$. We believe it is possible to implement the strategy outlined in Remark 21, and developed in the proof below, directly to $S(\mathbf{h})$ thanks to the representation of the evolution described in Remark 8. This would eliminate the need for conditions on $\mathbf{h}$ but it would make the proof below unnecessarily involved.

Proof of Lemma 20 A way to make the first step of the discussion in Remark 21 rigorous is to *coarse grain*, that is to approximate each h_N by a simple function obtained by averaging it over the element of a partition of $\mathbb{R}^N$ made by rectangles obtained as the Cartesian product of a finite number of measurable set of $\mathbb{R}$.

More precisely, we call $\mathscr{B} = \{B_k\}_{k=1}^K$ a *(measurable) partition* of $\mathbb{R}^N$ if $B_k \subset \mathbb{R}^N$ are measurable and $\bigcup_k B_k = \mathbb{R}^N$ while $B_k \cap B_{k'} = \emptyset$ if $k \neq k'$. Given a measurable partition $\mathcal{B}$ let $I_k(\underline{v}_N, \underline{w}_N)$ be the indicator function of $B_k \times B_k \subset \mathbb{R}^{2N}$ and define the *coarse graining* kernel:

$$C_{\mathscr{B}}(\underline{v}_N, \underline{w}_N) = \sum_{k=1}^{K} \frac{1}{\omega_k} I_k(\underline{v}_N, \underline{w}_N) \quad \text{with} \quad \omega_k = \int_{B_k} \gamma_N(\underline{v}_N) d\underline{v}_N \,.$$

Clearly, for every $\underline{w}_N$ we have

$$\int_{\mathbb{R}^N} C_{\mathcal{B}}(\underline{v}_N, \underline{w}_N)\gamma(\underline{v}_N) d\underline{v}_N = 1$$

while $C_{\mathcal{B}}(\underline{v}_n, \underline{w}_N) = C_{\mathcal{B}}(\underline{w}_N, \underline{v}_N)$. Given a function h_N is $L^1(\mathbb{R}^N)$ we can define its *coarse grained* version as

$$h_{N,\mathcal{B}}(\underline{v}_N) = \int_{\mathbb{R}^N} C_{\mathcal{B}}(\underline{v}_N, \underline{w}_N) h_N(\underline{w}_N)\gamma(\underline{w}_N) d\underline{w}_N \,.$$

Observe that, if $\underline{v}_N \in B_k$ then

$$h_{N,\mathcal{B}}(\underline{v}_N) = \frac{1}{\omega_k} \int_{B_k} \gamma(\underline{w}_N) h_N(\underline{w}_N) d\underline{w}_N \,.$$

This means that $h_{N,\mathcal{B}}(\underline{v}_N)$ is a simple function that assumes only K possible values. Finally we have $\int_{\mathbb{R}^N} h_{N,\mathcal{B}}(\underline{v}_N)\gamma(\underline{v}_N) d\underline{v}_N = \int_{\mathbb{R}^N} h_N(\underline{v}_N)\gamma(\underline{v}_N) d\underline{v}_N$.

Given measurable partitions $\mathcal{B} = \{B_k\}_{k=1}^K$ and $\mathcal{B}' = \{B'_j\}_{j=1}^J$of $\mathbb{R}^N$ and $\mathbb{R}^M$ respectively, we can define the *product* partition $\mathcal{B} \times \mathcal{B}' = \{B_k \times B'_j \,|k = 1, \ldots, K \quad j = 1, \ldots, J\}$ of $\mathbb{R}^{N+M}$. Observe that the coarse graining kernel of $\mathcal{B} \times \mathcal{B}'$ satisfies

$$C_{\mathcal{B}\times\mathcal{B}'}(\underline{v}_N, \underline{v}'_M, \underline{w}_N, \underline{w}'_M) = C_{\mathcal{B}}(\underline{v}_N, \underline{w}_N) C_{\mathcal{B}'}(\underline{v}'_M, \underline{w}'_M) \,.$$

Finally, given a partition $\mathcal{B} = \{B_k\}_{k=1}^K$ of $\mathbb{R}$, and $\underline{k} = (k_1, \ldots, k_N) \in \{1, \ldots, K\}^N$ we consider the set $B_{\underline{k}} = \times_i B_{k_i} \subset \mathbb{R}^N$. Clearly the $B_{\underline{k}}$ form a measurable partition of $\mathbb{R}^N$ that we will denote as $\mathcal{B}^N$. As before, we can define the coarse graining kernel for $\mathcal{B}^N$ as

$$C_{\mathcal{B}^N}(\underline{v}_N, \underline{w}_N) = \sum_{\underline{k}\in\{1,\ldots,K\}^N} \frac{1}{\omega_{\underline{k}}} I_{\underline{k}}(\underline{v}_N, \underline{w}_N)$$

where $\omega_{\underline{k}} = \prod_{i=1}^{N} \omega_{k_i}$ and $I_{\underline{k}}(\underline{v}_N, \underline{w}_N)$ is the characteristic function of $B_{\underline{k}} \times B_{\underline{k}} \in \mathbb{R}^{2N}$. Moreover the coarse grained version of $h_N \in L^1(\mathbb{R}^N, \gamma_N)$ is

$$h_{N,\mathcal{B}^N}(\underline{v}_N) = \int_{\mathbb{R}^N} \gamma(\underline{w}_N) C_{\mathcal{B}^N}(\underline{v}_N, \underline{w}_N) h_N(\underline{w}_N) d\underline{w}_N \,.$$

Again, if $\underline{v}_N \in B_{\underline{k}}$ we have

$$h_{N,\mathcal{B}^N}(\underline{v}_N) = \frac{1}{\omega_{\underline{k}}} \int_{B_{\underline{k}}} h_N(\underline{v}_N)\gamma_N(\underline{v}_N) d\underline{v}_N := \bar{h}_{N,\mathcal{B}^N}(\underline{k})$$

and $h_{N,\mathcal{B}^N}(\underline{v}_N)$ assumes only the K^N possible values $\bar{h}_{N,\mathcal{B}^N}(\underline{k})$. Observe finally that, since

$$C_{\mathcal{B}^N}(\underline{v}_N, \underline{w}_N) = \prod_{i=1}^{N} C_{\mathcal{B}}(v_i, w_i)\,,$$

we can write

$$h_{N-1,\mathcal{B}^{N-1}}(\underline{v}_{N-1}) = \int_{\mathbb{R}^N} \gamma(\underline{w}_N) C_{\mathcal{B}^N}(\underline{v}_N, \underline{w}_N) h_{N-1}(\underline{w}_{N-1}) d\underline{w}_N \,. \tag{76}$$

Given a state $\mathbf{h}$ and a partition $\mathcal{B}$ of $\mathbb{R}^N$, we define the coarse grained version $\mathbf{h}_{\mathcal{B}}$ of $\mathbf{h}$ over $\mathcal{B}$ by setting $h_{\mathcal{B},N} = h_{N,\mathcal{B}^N}$. Since $x\log(x)$ is convex in x and $(x-y)(\log(x) - \log(y))$ is jointly convex in x and y, for every partition $\mathcal{B}$ of $\mathbb{R}$, we get

$$S(\mathbf{h}_{\mathcal{B}}) \le S(\mathbf{h}), \qquad \Psi(\mathbf{h}_{\mathcal{B}}) \le \Psi(\mathbf{h}), \qquad E(\mathbf{h}_{\mathcal{B}}) = E(\mathbf{h}) \tag{77}$$

where in the inequality for Ψ we used (76). On the other hand, we have the following Lemma.

Lemma 23 *Given* $\mathbf{h}$*, for every* ϵ *we can find a finite measurable partition* $\mathcal{B}$ *of* $\mathbb{R}$ *such that*

$$S(\mathbf{h}) - S(\mathbf{h}_{\mathcal{B}}) \le \epsilon$$

Proof See Appendix A.4. □

We thus claim that to prove Lemma 20 we just need to show that, for every finite partition $\mathcal{B}$ of $\mathbb{R}$ and every state $\mathbf{h}$ we have

$$S(\mathbf{h}_{\mathcal{B}}) \le E(\mathbf{h}_{\mathcal{B}}) \log E(\mathbf{h}_{\mathcal{B}}) + \frac{\mu}{\rho} \Psi(\mathbf{h}_{\mathcal{B}})\,. \tag{78}$$

To see this observe that Lemma 23, together with (77) and (78), implies that for every ϵ we can find a partition $\mathcal{B}$ such that

$$\begin{aligned} S(\mathbf{h}) \le S(\mathbf{h}_{\mathcal{B}}) + \epsilon &\le E(\mathbf{h}_{\mathcal{B}}) \log E(\mathbf{h}_{\mathcal{B}}) + \frac{\mu}{\rho}\Psi(\mathbf{h}_{\mathcal{B}}) + \epsilon \\ &\le E(\mathbf{h}) \log E(\mathbf{h}) + \frac{\mu}{\rho}\Psi(\mathbf{h}) + \epsilon\,. \end{aligned}$$

Thus we consider a given finite partition $\mathcal{B} = \{B_k\}_{k=1}^K$ and a given state $\mathbf{h}$. Since $h_{\mathcal{B},N}$ takes only finitely many values, it should be possible to transform the integrals defining $E(\mathbf{h}_{\mathcal{B}})$, $S(\mathbf{h}_{\mathcal{B}})$ and $\Psi(\mathbf{h}_{\mathcal{B}})$ into summations. To do this, given $\underline{k} \in \{1, \ldots, K\}^N$, we define the *occupation numbers* $\underline{n}(\underline{k}) = (n_1(\underline{k}), \ldots, n_K(\underline{k})) \in \mathbb{N}^K$ as

$$n_q(\underline{k}) = \sum_i \delta_{q,k_i}\,.$$

That is $n_q(\underline{k})$is the number of i such that $k_i = q$. In other words, if $\underline{v}_N \in B_{\underline{k}}$ then there are $n_q(\underline{k})$ particles with velocity in B_q.

The fact that h_N is invariant under permutation of its arguments implies that $\bar{h}_{N,\mathcal{B}^N}(\underline{k})$ depends only on $\underline{n}(\underline{k})$ or, more precisely, if $\underline{n}(\underline{k}) = \underline{n}(\underline{k}')$ then $\bar{h}_{N,\mathcal{B}^N}(\underline{k}) = \bar{h}_{N,\mathcal{B}^N}(\underline{k}')$. This allow us to define the function $F : \mathbb{N}^K \to \mathbb{R}$ given by

$$F(\underline{n}) = \bar{h}_N(\underline{k}) \quad \text{if} \quad \underline{n} = \underline{n}(\underline{k}), \quad \text{and} \quad N = \sum_{k=1}^{K} n_k := |\underline{n}| \,.$$

Using this definition and the fact that $\sum_{k=1}^{K} \omega_k = 1$, we can now write

$$\begin{aligned} E(\mathbf{h}_{\mathcal{B}}) &= \sum_N a_N \sum_{\underline{k}\in\{1,\ldots,K\}^N} \bar{h}_{N,\mathcal{B}^N}(\underline{k})\omega_{\underline{k}} \\ &= \sum_N \frac{e^{-\frac{\mu}{\rho}}}{N!} \left(\frac{\mu}{\rho}\right)^N \sum_{|\underline{n}|=N} \binom{N}{n_1,\ldots,n_K} F(\underline{n}) \prod_{k=1}^{K} \omega_k^{n_k} \\ &= \sum_{\underline{n}\in\mathbb{N}^K} F(\underline{n}) \prod_{k=1}^{K} \pi_{\alpha_k}(n_k) := \widetilde{E}_{\underline{\alpha}_K}(F) \end{aligned} \tag{79}$$

where $\underline{\alpha}_K = (\alpha_1, \ldots, \alpha_K)$ with $\alpha_k = \mu\omega_k/\rho$ and

$$\pi_\alpha(n) = e^{-\alpha} \frac{\alpha^n}{n!} \,,$$

that is π_{α_k} is the Poisson distribution with expected value α_k. Similarly we have

$$\begin{aligned} S(\mathbf{h}_{\mathcal{B}}) &= \sum_N a_N \sum_{\underline{k}\in\{1,\ldots,K\}^N} \bar{h}_{N,\mathcal{B}^N}(\underline{k}) \log(\bar{h}_{N,\mathcal{B}^N}(\underline{k}))\omega_{\underline{k}} \\ &= \sum_{\underline{n}\in\mathbb{N}^K} F(\underline{n}) \log(F(\underline{n})) \prod_{k=1}^{K} \pi_{\alpha_k}(n_k) := \widetilde{S}_{\underline{\alpha}_K}(F) \end{aligned} \tag{80}$$

Finally setting $\underline{n}^q = (n_1, \ldots, n_q + 1, \ldots, n_K)$ we get

$$\begin{aligned} \Psi(\mathbf{h}_{\mathcal{B}}) = \sum_N a_N \sum_{\underline{k}\in\{1,\ldots,K\}^N} \sum_{q=1}^{K} &(\bar{h}_{N+1,\mathcal{B}^{N+1}}(\underline{k}, q) - \bar{h}_{N,\mathcal{B}^N}(\underline{k}))\cdot \\ &\times (\log \bar{h}_{N+1,\mathcal{B}^{N+1}}(\underline{k}, q) - \log \bar{h}_{N,\mathcal{B}^N}(\underline{k}))\omega_{\underline{k}}\omega_q \\ &= \frac{\rho}{\mu} \sum_{q=1}^{K} \alpha_q \sum_{\underline{n}\in\mathbb{N}^K} \left(F(\underline{n}^q) - F(\underline{n})\right) \left(\log F(\underline{n}^q) - \log F(\underline{n})\right) \prod_{k=1}^{K} \pi_{\alpha_k}(n_k) \\ &:= \frac{\rho}{\mu} \widetilde{\Psi}_{\underline{\alpha}_K}(F) \,. \end{aligned} \tag{81}$$

so that, to prove (78), we need to show that, for every $F : \mathbb{N}^K \to \mathbb{R}_+$ and for every K and $\underline{\alpha}_K \in \mathbb{R}_+^K$, if $\widetilde{S}_K(F) < \infty$ then

$$\widetilde{S}_{\underline{\alpha}_K}(F) \le \widetilde{\Psi}_{\underline{\alpha}_K}(F) + \widetilde{E}_{\underline{\alpha}_K}(F) \log \widetilde{E}_{\underline{\alpha}_K}(F) \,. \tag{82}$$

We will prove (82) by induction over K. Assume that (82) is valid for every index less than K for some $K > 1$ and write

$$\widetilde{S}_{\underline{\alpha}_{K-1}}(F(\cdot, n_K)) = \sum_{\underline{n}' \in \mathbb{N}^{K-1}} F(\underline{n}', n_K) \log F(\underline{n}', n_K) \prod_{k=1}^{K-1} \pi_{\alpha_k}(n_k)$$

and similar expression for $E_{\underline{\alpha}_{K-1}}(F(\cdot, n_K))$ and $\Psi_{\underline{\alpha}_{K-1}}(F(\cdot, n_K))$.

Using the inductive hypothesis we obtain

$$\widetilde{S}_{\underline{\alpha}_K}(F) = \sum_{n_K=0}^{\infty} \widetilde{S}_{\underline{\alpha}_{K-1}}(F(\cdot, n_K))\pi_{\alpha_K}(n_K) \leq \sum_{n_K=0}^{\infty} \widetilde{\Psi}_{\underline{\alpha}_{K-1}}(F(\cdot, n_K))\pi_{\alpha_K}(n_K)$$
$$+ \sum_{n_K=0}^{\infty} \widetilde{E}_{\underline{\alpha}_{K-1}}(F(\cdot, n_K)) \log \widetilde{E}_{\underline{\alpha}_{K-1}}(F(\cdot, n_K))\pi_{\alpha_K}(n_K)\,.$$

Calling $F_1(n_K) = \widetilde{E}_{\underline{\alpha}_{K-1}}(F(\cdot, n_K))$ and using the inductive hypothesis again we get

$$\sum_{n_K=0}^{\infty} \widetilde{E}_{\underline{\alpha}_{K-1}}(F(\cdot, n_K)) \log \widetilde{E}_{\underline{\alpha}_{K-1}}(F(\cdot, n_K))\pi_{\alpha_K}(n_K) = \widetilde{S}_{\alpha_K}(F_1)$$
$$\leq \widetilde{\Psi}_{\alpha_K}(F_1) + \widetilde{E}_{\alpha_K}(F_1) \log \widetilde{E}_{\alpha_K}(F_1)$$

so that

$$\widetilde{S}_{\underline{\alpha}_K}(F) \leq \sum_{n_K=0}^{\infty} \widetilde{\Psi}_{\underline{\alpha}_{K-1}}(F(\cdot, n_K))\pi_{\alpha_K}(n_K) + \widetilde{\Psi}_{\alpha_K}(F_1) + \widetilde{E}_{\alpha_K}(F_1) \log \widetilde{E}_{\alpha_K}(F_1)\,. \quad (83)$$

Observing that $\widetilde{E}_{\alpha_K}(F_1) = \widetilde{E}_{\underline{\alpha}_K}(F)$ and that, by convexity,

$$\widetilde{\Psi}_{\alpha_K}(F_1) = \alpha_K \sum_{n=0}^{\infty} (F_1(n+1) - F_1(n))(\log F_1(n+1) - \log F_1(n))\pi_{\alpha_K}(n)$$
$$\leq \alpha_K \sum_{\underline{n} \in \mathbb{N}^K} \left(F(\underline{n}^K) - F(\underline{n})\right) \left(\log F(\underline{n}^K) - \log F(\underline{n})\right) \prod_{k=1}^{K} \pi_{\alpha_k}(n_k)$$

we get (82) for K. Thus, by induction, to prove (82) for every K we just need to prove it for $K = 1$. This is the content of the following Lemma.

Lemma 24 *Let π_α be the Poisson distribution on $\mathbb{N}$ with expected value $\alpha > 0$ and $f : \mathbb{N} \to \mathbb{R}^+$ be such that*

$$\sum_{n=0}^{\infty} f(n) \log f(n)\pi_\alpha(n) < \infty\,,$$

then we have

$$\sum_{n=0}^{\infty} f(n) \log f(n)\pi_\alpha(n) \leq \left(\sum_{n=0}^{\infty} f(n)\pi_\alpha(n)\right) \log \left(\sum_{n=0}^{\infty} f(n)\pi_\alpha(n)\right)$$
$$+ \alpha \sum_{n=0}^{\infty} (f(n+1) - f(n)) (\log f(n+1) - \log f(n))\, \pi_\alpha(n)\,. \quad (84)$$

Proof Observe first that since $\alpha\pi_\alpha(n) = (n+1)\pi_\alpha(n+1)$ we get

$$\begin{aligned} &\alpha \sum_{n=0}^{\infty} (f(n+1) - f(n)) (\log f(n+1) - \log f(n)) \pi_\alpha(n) \\ &\quad = \sum_{n=1}^{\infty} n\,(f(n) - f(n-1)) (\log f(n) - \log f(n-1)) \pi_\alpha(n)\,. \end{aligned}$$

Let now $\pi_{\alpha,N}(n)$ be the binomial distribution with parameters N and α/N, that is

$$\pi_{\alpha,N}(n) = \binom{N}{n} \left(\frac{\alpha}{N}\right)^n \left(1 - \frac{\alpha}{N}\right)^{N-n}\,.$$

We will prove by induction that for every N and every $\alpha \le N$ we have

$$\begin{aligned} \sum_{n=0}^{N} f(n) \log f(n) \pi_{\alpha,N}(n) \le &\left(\sum_{n=0}^{N} f(n)\pi_{\alpha,N}(n)\right) \log \left(\sum_{n=0}^{N} f(n)\pi_{\alpha,N}(n)\right) \\ &+ \sum_{n=1}^{N} n\,(f(n) - f(n-1)) (\log f(n) - \log f(n-1)) \pi_{\alpha,N}(n) \end{aligned} \tag{85}$$

so that, taking the limit for $N \to \infty$, we will obtain (84). The base case $N = 1$ is covered by the following Lemma. □

Lemma 25 *Let $\mu_x \ge 0$, $x \in \{0,1\}$, be such that $\mu_0 + \mu_1 = 1$ then for every function $f : \{0,1\} \to \mathbb{R}^+$ we have*

$$\begin{aligned} \sum_{x=0,1} f(x) \log f(x)\mu_x \le &\left(\sum_{x=0,1} f(x)\mu_x\right) \log \left(\sum_{x=0,1} f(x)\mu_x\right)' \\ &+ \mu_0\mu_1 (f(1) - f(0)) (\log f(1) - \log f(0))\,. \end{aligned} \tag{86}$$

Proof Calling $h(0) = f(0)/(\mu_0 f(0) + \mu_1 f(1))$ and $h(1) = f(1)/(\mu_0 f(0) + \mu_1 f(1))$, (86) becomes

$$\sum_{x=0,1} h(x) \log h(x)\mu_x \le \mu_0\mu_1 (h(1) - h(0)) (\log h(1) - \log h(0))\,.$$

Since $\mu_0 h(0) + \mu_1 h(1) = 1$ we can write $h(0) = 1 + \delta\mu_1$ and $h(1) = 1 - \delta\mu_0$ and we get

$$\begin{aligned} &\sum_{x=0,1} h(x) \log h(x)\mu_x \\ &\quad = \mu_0\mu_1\delta(\log(1+\delta\mu_1) - \log(1-\delta\mu_0)) + \mu_0 \log(1+\delta\mu_1) + \mu_1 \log(1-\delta\mu_0) \\ &\quad \le \mu_0\mu_1\delta(\log(1+\delta\mu_1) - \log(1-\delta\mu_0)) \\ &\quad = \mu_0\mu_1 (h(1) - h(0)) (\log h(1) - \log h(0)) \end{aligned}$$

where we have used concavity of the logarithm. □

Assume now that (85) holds for every index less than N. Given $\alpha \le N$ call $\beta = (N-1)\alpha/N$ so that $\beta \le N-1$. Define also $\mu_0 = 1 - \alpha/N$, $\mu_1 = \alpha/N$, and observe that, for every $J : \mathbb{N} \to \mathbb{R}$,

$$\sum_{n=0}^{N} J(n)\pi_{\alpha,N}(n) = \sum_{x=0,1} \sum_{n=0}^{N-1} J(n+x)\pi_{\beta,N-1}(n)\mu_x \,. \tag{87}$$

Calling

$$\bar{f}(x) = \sum_{n=0}^{N-1} f(n+x)\pi_{\beta,N-1}(n)$$

and using (87) and the inductive hypothesis for index $N-1$, we get

$$\begin{aligned}
\sum_{n=0}^{N} f(n)\log f(n)\pi_{\alpha,N}(n) &= \sum_{x=0,1} \sum_{n=0}^{N-1} f(n+x)\log f(n+x)\pi_{\beta,N-1}(n)\mu_x \\
&\le \sum_{x=0,1} \bar{f}(x)\log \bar{f}(x)\mu_x \\
&\quad + \sum_{x=0,1} \sum_{n=1}^{N-1} n\left(f(n+x) - f(n-1+x)\right) \\
&\quad \cdot \left(\log f(n+x) - \log f(n-1+x)\right)\pi_{\beta,N-1}(n)\mu_x
\end{aligned}$$

while using Lemma 25 for the first term in the second line delivers

$$\begin{aligned}
\sum_{n=0}^{N} f(n)\log f(n)\pi_{\alpha,N}(n) &\le \left(\sum_{n=0}^{N} f(n)\pi_{\alpha,N}(n)\right)\log\left(\sum_{n=0}^{N} f(n)\pi_{\alpha,N}(n)\right) \\
&\quad + \mu_0\mu_1(\bar{f}(1) - \bar{f}(0))(\log \bar{f}(1) - \log \bar{f}(0)) \\
&\quad + \sum_{x=0,1} \sum_{n=1}^{N-1} n\left(f(n+x) - f(n-1+x)\right) \\
&\quad \cdot \left(\log f(n+x) - \log f(n-1+x)\right)\pi_{\beta,N-1}(n)\mu_x
\end{aligned} \tag{88}$$

Finally using the joint convexity in (x, y) of the function $(x - y)(\log x - \log y)$ and the fact that $\mu_0 < 1$ we can write

$$\begin{aligned}
&\mu_0\mu_1(\bar{f}(1) - \bar{f}(0))(\log \bar{f}(1) - \log \bar{f}(0)) \\
&\quad \le \mu_1 \sum_{n=0}^{N-1} (f(n+1) - f(n))(\log f(n+1) - \log f(n))\pi_{\beta,N-1}(n)
\end{aligned}$$

that inserted in (88) gives

$$\begin{aligned}
\sum_{n=0}^{N} f(n)\log f(n)\pi_{\alpha,N}(n) &\le \left(\sum_{n=0}^{N} f(n)\pi_{\alpha,N}(n)\right)\log\left(\sum_{n=0}^{N} f(n)\pi_{\alpha,N}(n)\right) \\
&\quad + \sum_{x=0,1} \sum_{n=1}^{N-1} (n+x)\left(f(n+x) - f(n-1+x)\right) \\
&\quad \cdot \left(\log f(n+x) - \log f(n-1+x)\right)\pi_{\beta,N-1}(n)\mu_x \,.
\end{aligned}$$

Changing summation variables from (x, n) to $(x, n + x)$ and using (87) we obtain (85) for index N. Thus (85) is valid for every $N \geq 1$ and every $\alpha \leq N$.

To complete the proof of Lemma 24 we need to show that we can take the limit for $N \to \infty$ in (85). To this end observe that given α, for N large enough we have $0 < \left(1 - \frac{\alpha}{N}\right)^N \leq 2e^{-\alpha}$. Thus for large N and $\alpha < n \leq N$ we get

$$\begin{aligned} \pi_{\alpha,N}(n) &\leq 2e^{-\alpha}\frac{(\alpha)^n}{n!}\left(1 - \frac{\alpha}{N}\right)^{-n}\prod_{i=1}^{n}\left(1 - \frac{i}{N}\right) \\ &\leq 2e^{-\alpha}\frac{(\alpha)^n}{n!}\left(1 - \frac{\alpha}{N}\right)^{-\lfloor\alpha\rfloor}\prod_{i=1}^{\lfloor\alpha\rfloor}\left(1 - \frac{i}{N}\right) \leq 4\pi_\alpha(n)\,. \end{aligned} \tag{89}$$

Using Dominated Convergence, (89) implies that, if $f(n)$ is bounded below and $\sum_{n=0}^{\infty} f(n)$ $\pi_\alpha(n) \leq \infty$ then

$$\lim_{N\to\infty}\sum_{n=0}^{N} f(n)\pi_{\alpha,N}(n) = \sum_{n=0}^{\infty} f(n)\pi_\alpha(n)\,.$$

We can now let $N \to \infty$ in (85) to obtain (84). This concludes the proof of Lemma 24. □

To sum up, the validity of (84) together with the inductive argument in (83) shows that (82) is valid for every K and $\alpha_k, k = 1, \ldots, K$. This in turn, together with (79), (80) and (81), establishes the validity of (78) for every state $\mathbf{h}$ and every partition $\mathcal{B}$ of $\mathbb{R}$. This, together with Lemma 24 completes the proof of Lemma 20. □

Observe now that if $\mathbf{f}$ is a probability distribution $E(\mathbf{h}) = 1$ so that Lemma 20, together with Lemma 19, gives

$$\frac{d_+}{dt}S(\mathbf{h}(t)) \leq -\rho S(\mathbf{h}(t)) \tag{90}$$

To complete the proof of Theorem 4 we have to show that (90) implies (15). To this end, take $\rho' < \rho$, assume that there exists t such that $S(\mathbf{h}(t)) > e^{-\rho' t}S(\mathbf{h}(0))$ and let

$$T = \inf\{t \geq 0 \,|\, S(\mathbf{h}(t)) > e^{-\rho' t}S(\mathbf{h}(0))\}\,.$$

By continuity we get $S(\mathbf{h}(T)) = e^{-\rho' T}S(\mathbf{h}(0))$. From (90), for every ϵ we can find δ such that

$$S(\mathbf{h}(T + h)) \leq (1 - \rho h)e^{-T\rho'}S(\mathbf{h}(0)) + h\epsilon$$

for every $h \leq \delta$. Choosing $\epsilon = (\rho - \rho')e^{-T\rho'}S(\mathbf{h}(0))$ we get

$$S(\mathbf{h}(T + h)) \leq e^{-(T+h)\rho'}S(\mathbf{h}(0))$$

which implies that $S(\mathbf{h}(t)) \leq e^{-t\rho'}S(\mathbf{h}(0))$ for every $t \geq 0$ and every $\rho' < \rho$. □

3.5 Derivation of (17)

To prove (17), we observe that $\eta(t)$ and $g(v, t)$ in (18) satisfy the equations

$$\begin{aligned} \dot\eta(t) &= \mu - \rho\eta(t) \\ \dot g(v, t) &= \frac{\mu}{\eta(t)}(\gamma(v) - g(v, t))\,. \end{aligned}$$

Setting $\mathbf{f}(t) = (f_0(t), f_1(t), f_2(t), \ldots)$ with

$$f_N(\underline{v}_N, t) = e^{-\eta(t)} \frac{\eta(t)^N}{N!} \prod_{i=1}^{N} g(v_i, t)$$

we get

$$\begin{aligned}
\frac{d}{dt} f_N(\underline{v}_N, t) &= (\mu - \rho\eta(t)) e^{-\eta(t)} \frac{\eta(t)^{N-1}}{(N-1)!} \left(1 - \frac{\eta(t)}{N}\right) \prod_{i=1}^{N} g(v_i, t) \\
&\quad + \mu e^{-\eta(t)} \frac{\eta(t)^{N-1}}{N!} \sum_i \left((\gamma(v_i) - g(v_i, t)) \prod_{j \neq i} g(v_j, t) \right) \\
&= \rho e^{-\eta(t)} \frac{\eta(t)^{N+1}}{N!} \prod_{i=1}^{N} g(v_i, t) - \rho e^{-\eta(t)} \frac{\eta(t)^N}{(N-1)!} \prod_{i=1}^{N} g(v_i, t) \\
&\quad + \mu e^{-\eta(t)} \frac{\eta(t)^{N-1}}{N!} \sum_i \gamma(v_i) \prod_{j \neq i} g(v_j, t) - \mu e^{-\eta(t)} \frac{\eta(t)^N}{N!} \prod_{i=1}^{N} g(v_i, t) \\
&= \rho((\mathcal{O}\mathbf{f}(t))_N(\underline{v}) - N f_N(\underline{v}_N, t)) + \mu((\mathcal{I}\mathbf{f}(t))_N(\underline{v}) - f_N(\underline{v}_N, t)) .
\end{aligned}$$

Thus $\mathbf{f}(t)$ solves (1) with $\tilde{\lambda} = 0$. Clearly $\mathbf{f}(t) \in D^1$ for every $t \geq 0$ so that, by Remark 7, $\mathbf{f}(t) = e^{t\mathcal{T}} \mathbf{f}(0)$.

3.6 Proof of Theorem 5

Given a continuous and bounded test function $\phi_k : \mathbb{R}^k \to \mathbb{R}$, symmetric with respect to the permutation of its variables, we define

$$(\mathbf{f}_n, \phi_k)_{k,n} = \left(\frac{\rho}{\mu_n}\right)^k \sum_{N \geq k} \frac{N!}{(N-k)!} \int_{\mathbb{R}^N} f_{n,N}(\underline{v}_N) \phi_k(\underline{v}_k) d\underline{v}_N .$$

What we need to show is that, if $\mathbf{f}_n$ forms a chaotic sequence and $\phi : \mathbb{R} \to \mathbb{R}$ is a test function then

$$\lim_{n\to\infty} (e^{\mathcal{L}_n t} \mathbf{f}_n, \phi^{\otimes k})_{k,n} = \left(\lim_{n\to\infty} (e^{\mathcal{L}_n t} \mathbf{f}_n, \phi)_{1,n} \right)^k$$

which implies propagation of chaos.

The argument to prove propagation of chaos introduced in [16] is based on the power series expansion of $e^{\lambda K_N t}$, which converges since K_N is a bounded operator. After this, one can exploit a cancellation between Q_N and $\binom{N}{2}\mathrm{Id}$, see (5), when they act on a function ϕ_k depending only on $k < N$ variables, see Sect. 3 of [16]. In the present case the analogue of such an argument formally works but it cannot be applied directly since, being $\mathcal{K}$ unbounded, the power series expansion of $e^{\tilde{\lambda}_n \mathcal{K} t}$ does not converge. To avoid this problem, one may try to use the convergent expansion (27) introduced in Sect. 3.1. But the different treatment of Q_N and $\binom{N}{2}\mathrm{Id}$ in (27) would make it very hard to see the needed cancellation.

Thus we will introduce a partial expansion of $e^{\tilde{\lambda}_n \mathcal{K} t}$ and combine it with (33) and (40). The idea is to expand this exponential in the *least possible way* to exploit the central cancellations of McKean's argument. We first decompose K_N as

$$K_N = K_k + \widetilde{K}_{N-k} + (N-k) G_k$$

with

$$\widetilde{K}_{N-k} = \sum_{k+1 \le i < j \le N} (R_{i,j} - \mathrm{Id})$$

$$G_k = \frac{1}{N-k} \sum_{i=1}^{k} \sum_{j=k+1}^{N} (R_{i,j} - \mathrm{Id})$$

and obtain

$$e^{\tilde{\lambda}_n K_N t} \phi_k = e^{\tilde{\lambda}_n K_k t} \phi_k + (N-k) \tilde{\lambda}_n \int_0^t e^{\tilde{\lambda}_n K_N (t-s)} G_k e^{\tilde{\lambda}_n K_k s} \phi_k ds \tag{91}$$

where we used that K_N is a bounded operator on $C^0(\mathbb{R}^N)$ and that $\widetilde{K}_{N-k}\phi_k = 0$. Since we are interested in integrating (91) against a symmetric function f_N we can write

$$G_k[\phi_k](\underline{v}_{k+1}) = \sum_{i=1}^{k} \int \frac{d\theta}{2\pi} [\phi(v_1, \dots, v_{i-1}, v_i \cos\theta + v_{k+1} \sin\theta, v_{i+1}, \dots, v_k) - \phi(\underline{v}_k)]$$

To iterate we need to apply (91) to the factor $e^{\tilde{\lambda}_n K_N (t-s)}$ inside the integral in (91) itself. Since $G_k e^{\tilde{\lambda}_n K_k s} \phi_k$ is a function of $k+1$ variables we now have to write

$$K_N = K_{k+1} + \widetilde{K}_{N-k-1} + (N-k-1) G_{k+1} \,.$$

Iterating this procedure we get

$$\begin{aligned} e^{\tilde{\lambda}_n K_N t} \phi_k &= e^{\tilde{\lambda}_n K_k t} \phi_k \\ &\quad + \sum_{p=1}^{N-k} \frac{\tilde{\lambda}_n^p (N-k)!}{(N-k-p)!} \int_{0<t_1<\cdots<t_p<t} e^{\tilde{\lambda}_n K_{k+p}(t-t_p)} G_{k+p-1} e^{\tilde{\lambda}_n K_{k+p-1}(t_p - t_{p-1})} \\ &\quad \cdots e^{(t_2-t_1)\tilde{\lambda}_n K_{k+1}} G_k e^{\tilde{\lambda}_n K_k t_1} \phi_k \, dt_p \cdots dt_1 \end{aligned}$$

so that

$$\begin{aligned} (e^{\tilde{\lambda}_n \mathcal{K} t} \mathbf{f}_n, \phi_k)_{k,n} &= (\mathbf{f}_n, e^{\tilde{\lambda}_n K_k t} \phi_k)_{k,n} \\ &\quad + \sum_{p=1}^{\infty} \lambda^p \int_{0<t_1<\cdots<t_p<t} \Big(\mathbf{f}_n, e^{\tilde{\lambda}_n K_{k+p}(t-t_p)} G_{k+p-1} e^{\tilde{\lambda}_n K_{k+p-1}(t_p - t_{p-1})} \\ &\quad \cdots e^{(t_2-t_1)\tilde{\lambda}_n K_{k+1}} G_k e^{\tilde{\lambda}_n K_k t} \phi_k \Big)_{k+p,n} dt_p \cdots dt_1 \end{aligned} \tag{92}$$

where the factor λ^p in the second line of (92), comes from (25) and (19).

Observe now that the $R_{i,j}$ are averaging operators so that $\|R_{i,j}\|_\infty \le 1$ which gives

$$\left\| e^{t \tilde{\lambda}_n K_N} \right\|_\infty = e^{-t \tilde{\lambda} \binom{N}{2}} \left\| e^{t \tilde{\lambda} \sum_{1 \le i < j \le N} R_{i,j}} \right\|_\infty \le 1 \,.$$

For the same reason we have

$$\|G_k\|_\infty \le \frac{1}{N-k} \sum_{i=1}^{k} \sum_{j=k+1}^{N} (\|R_{i,j}\|_\infty + 1) \le 2k \,.$$

Using (21) we get

$$\left| (e^{\tilde{\lambda}_n \mathcal{K} t} \mathbf{f}_n, \phi_k)_{k,n} \right| \le \left(\frac{\rho}{\mu_n} \right)^k \|\mathbf{f}_n\|_1^{(k)} \|\phi_k\|_\infty$$

$$+\sum_{p=1}^{\infty}\frac{\lambda^p t^p}{p!}\prod_{i=k}^{k+p-1}\|G_i\|_\infty\left(\frac{\rho}{\mu_n}\right)^{k+p}\|\mathbf{f}_n\|_1^{(k+p)}\|\phi_k\|_\infty$$
$$\le \|\phi_k\|_\infty K^k\sum_{p=0}^{\infty}2^p\lambda^p t^p K^p\binom{k+p-1}{p}.$$

Observe that the series in the last line converges for $\lambda K t < 1/2$. On the other hand, since $\lim_{n\to\infty}\tilde{\lambda}_n = 0$, for every t we have

$$\lim_{n\to\infty}(\mathbf{f}_n, e^{\tilde{\lambda}_n K_k t}\phi_k)_{k,n} = \lim_{n\to\infty}(\mathbf{f}_n, \phi_k)_{k,n}$$

and similarly, calling $G_k^{*p} = \prod_{i=0}^{p} G_{k+i}$,

$$\lim_{n\to\infty}\int_{0<t_1<\cdots<t_p<t}\Big(\mathbf{f}_n, e^{\tilde{\lambda}_n K_{k+p}(t-t_p)}G_{k+p-1}e^{\tilde{\lambda}_n K_{k+p-1}(t_p-t_{p-1})}$$
$$\cdots e^{(t_2-t_1)\tilde{\lambda}_n K_{k+1}}G_k e^{\tilde{\lambda}_n K_k t}\phi_k\Big)_{k+p,n}dt_p\cdots dt_1 = \lim_{n\to\infty}\left(\mathbf{f}_n, G_k^{*p}\phi_k\right)_{k+p,n}$$

so that we finally get

$$\lim_{n\to\infty}(e^{\tilde{\lambda}_n \mathcal{K}t}\mathbf{f}_n, \phi_k)_{k,n} = \lim_{n\to\infty}\sum_{p=0}^{\infty}\frac{\lambda^p t^p}{p!}\left(\mathbf{f}_n, G_k^{*p}\phi_k\right)_{k+p,n}. \tag{93}$$

Observe now that G_k acts as a derivation in the sense of [16], that is, for every ϕ_{k_1} and ψ_{k_2} with $k_1 + k_2 = k$, we have

$$G_k(\phi_{k_1}\otimes\psi_{k_2}) = (G_{k_1}\phi_{k_1})\otimes\psi_{k_2} + \phi_{k_1}\otimes(G_{k_2}\psi_{k_2}).$$

This implies that

$$\frac{1}{p!}G_k^{*p}(\phi_{k_1}\otimes\psi_{k_2}) = \sum_{p_1+p_2=p}\frac{1}{p_1!}\frac{1}{p_2!}(G_{k_1}^{*p_1}\phi_{k_1})\otimes(G_{k_2}^{*p_1}\psi_{k_2}). \tag{94}$$

Observing that if $\mathbf{f}_n$ forms a chaotic sequence then

$$\lim_{n\to\infty}(\mathbf{f}_n, \phi_{k_1}\otimes\psi_{k_2})_{k,n} = \lim_{n\to\infty}(\mathbf{f}_n, \phi_{k_1})_{k_1,n}\lim_{n\to\infty}(\mathbf{f}_n, \psi_{k_2})_{k_2,n} \tag{95}$$

we get

$$\lim_{n\to\infty}\sum_{p=0}^{\infty}\frac{\lambda^p t^p}{p!}(\mathbf{f}_n, G_k^{*p}\phi_{k_1}\otimes\psi_{k_2})_{k+p,n} \tag{96}$$
$$= \lim_{n\to\infty}\sum_{p_1=0}^{\infty}\frac{\lambda^{p_1}t^{p_1}}{p_1!}\left(\mathbf{f}_n, G_{k_1}^{*p_1}\phi_{k_1}\right)_{k_1+p_1,n}\lim_{n\to\infty}\sum_{p_2=0}^{\infty}\frac{\lambda^{p_2}t^{p_2}}{p_2!}\left(\mathbf{f}_n, G_{k_2}^{*p_2}\phi_{k_2}\right)_{k_2+p_2,n}$$

which implies that $e^{\tilde{\lambda}_n \mathcal{K}t}$ propagates chaos, at least for $t \le t_0 = \frac{1}{2\lambda K}$. Finally we need to verify that (21) still holds. Since $\mathbf{f}_n$ are positive $\|\mathbf{f}_n\|_1^{(r)} = N_r(\mathbf{f})$, see (43). Thus Corollary 9 implies that for every $t \ge 0$ we have $\|\mathbf{f}_n(t)\|^{(r)} \le K_1^r\left(\frac{\mu_n}{\rho}\right)^r$ with $K_1 = \max\{K, 1\}$. Thus $\mathbf{f}_n(t_0) = e^{\tilde{\lambda}_n \mathcal{K}t_0}\mathbf{f}_n$ forms a chaotic sequence that satisfies (21) with K_1 in place of K. Using $\mathbf{f}_n(t_0)$ as initial condition we get that propagation of chaos holds up to time $t_1 = \frac{1}{2\lambda K} + \frac{1}{2\lambda K_1}$. Iterating this argument we see that $e^{\tilde{\lambda}_n \mathcal{K}t}$ propagates chaos for every $t \ge 0$.

To add the *out* operator $\mathcal{O}$, we observe that from (93) we get

$$
\begin{aligned}
&\lim_{n\to\infty}(e^{(\tilde{\lambda}_n\mathcal{K}-\rho\mathcal{N})t}\mathbf{f}_n,\phi_k)_{k,n}\\
&\quad=\lim_{n\to\infty}\sum_{p=0}^{\infty}\frac{\lambda^p t^p}{p!}\left(\frac{\rho}{\mu}\right)^{k+p}\sum_{N\geq k+p}\frac{N!}{(N-k-p)!}e^{-\rho N t}\\
&\quad\quad\times\int f_{n,N}(\underline{v}_N)(G_k^{*p}\phi_k)(\underline{v}_{k+p})d\underline{v}_N\,.
\end{aligned}
\tag{97}
$$

Inserting (97) into (33), after some long algebra that we report in Appendix A.5, we obtain

$$
\lim_{n\to\infty}\left(e^{(\tilde{\lambda}_n\mathcal{K}+\rho(\mathcal{O}-\mathcal{N}))t}\mathbf{f}_n,\phi_k\right)_{k,n}=\lim_{n\to\infty}\sum_{p=0}^{\infty}\frac{t^p\lambda^p}{p!}\left(\mathbf{f}_n,e^{-\rho(k+p)t}G_k^{*p}\phi_k\right)_{k+p,n}\,.
\tag{98}
$$

It is not hard to see that (98) implies that $e^{(\tilde{\lambda}_n\mathcal{K}+\rho(\mathcal{O}-\mathcal{N}))t}$ propagates chaos.

Finally we consider the *in* operator $\mathcal{I}$. Observe that

$$
\begin{aligned}
\mu_n(\mathcal{I}\mathbf{f}_n,\phi_k)_{k,n}&=\mu_n\left(\frac{\rho}{\mu_n}\right)^k\sum_{N\geq k}\frac{(N-1)!}{(N-k)!}\int\sum_{i=1}^{N}f_{n,N-1}(\underline{v}_{N-1}^i)\gamma(v_i)\phi_k(\underline{v}_k)d\underline{v}_N\\
&=\mu_n\left(\frac{\rho}{\mu_n}\right)^k\sum_{N\geq k}\frac{(N-1)!}{(N-k)!}\int\Big((N-k)f_{n,N-1}(\underline{v}_{N-1})\gamma(v_N)\phi_k(\underline{v}_k)d\underline{v}_N\\
&\quad+kf_{n,N-1}(\underline{v}_{N-1})\phi_k(\underline{v}_{k-1},v_N)\gamma(v_N)d\underline{v}_N\Big)\\
&=\mu_n\left(\frac{\rho}{\mu_n}\right)^k\sum_{N>k}\frac{(N-1)!}{(N-1-k)!}\int f_{n,N-1}(\underline{v}_{N-1})\phi_k(\underline{v}_k)\gamma(v_N)d\underline{v}_{N-1}\\
&\quad+k\rho\left(\frac{\rho}{\mu_n}\right)^{k-1}\sum_{N\geq k-1}\frac{N!}{(N-(k-1))!}\\
&\quad\times\int f_{n,N}(\underline{v}_N)\phi_k(\underline{v}_{k-1},w)\gamma(v_N)d\underline{v}_N dw
\end{aligned}
$$

so that

$$
\mu_n(\mathcal{I}\mathbf{f}_n-\mathbf{f}_n,\phi_k)_{k,n}=(\mathbf{f}_n,I_k\phi_k)_{k-1,n}
\tag{99}
$$

where

$$
I_k[\phi_k](\underline{v}_{k-1}):=\rho k\int_{\mathbb{R}}\phi_k(\underline{v}_{k-1},w)e^{-\pi w^2}dw\,.
$$

which clearly act as a derivative in the sense of [16]. We can now use an expansion similar to (34)

$$
\begin{aligned}
e^{t\mathcal{L}_n}\mathbf{f}_n&=e^{(\tilde{\lambda}_n\mathcal{K}+\rho(\mathcal{O}-\mathcal{N}))t}\mathbf{f}_n\\
&\quad+\sum_{q=1}^{\infty}\mu_n^q\int_{0<t_1<\ldots<t_q<t}e^{(\tilde{\lambda}_n\mathcal{K}+\rho(\mathcal{O}-\mathcal{N}))(t-t_q)}(\mathcal{I}-\mathrm{Id})e^{(\tilde{\lambda}_n\mathcal{K}+\rho(\mathcal{O}-\mathcal{N}))(t_q-t_{q-1})}\\
&\quad\cdots(\mathcal{I}-\mathrm{Id})e^{(\tilde{\lambda}_n\mathcal{K}+\rho(\mathcal{O}-\mathcal{N}))t_1}\mathbf{f}_n\,dt_1\cdots dt_n
\end{aligned}
\tag{100}
$$

that combined (99) with (98) gives

$$\lim_{n\to\infty}(e^{\mathcal{L}_n t}\mathbf{f}_n, \phi_k)_{k,n} = \lim_{n\to\infty}\sum_{q\geq 0}\sum_{p_0,p_1,\ldots,p_q\geq 0}\rho^q\lambda^{|p|}e^{-\rho k t}$$
$$\cdot\int_{0\leq t_q\leq\cdots\leq t_1\leq t}\prod_{i=0}^{q}e^{-\rho(t_i-t_{i+1})(|p|_i-i)}\frac{(t_i-t_{i+1})^{p_i}}{p_i!}dt_1\cdots dt_q\cdot$$
$$\left(\mathbf{f}_n, G^{*p_q}_{k+|p|_q-q}I_{k+|p|_q-q+1}\cdots G^{*p_1}_{k+p_0-1}I_{k+p_0}G^{*p_0}_k\phi_k\right)_{k+|p|-q,n} \tag{101}$$

where $|p|_i=\sum_{j=0}^{i-1}p_j$ and $t_0=t$, $t_{q+1}=0$ and the order of the t_i in the integral is inverted due to the inversion of the order of the operators when taking the adjoint. From (101) it follows, after more long algebra reported in Appendix A.5, that, if $k_1+k_2=k$, then

$$\lim_{n\to\infty}(e^{\mathcal{L}_n t}\mathbf{f}_n, \phi_{k_1}\otimes\psi_{k_2})_{k,n} = \lim_{n\to\infty}(e^{\mathcal{L}_n t}\mathbf{f}_n, \phi_{k_1})_{k_1,n}\lim_{n\to\infty}(e^{\mathcal{L}_n t}\mathbf{f}_n, \psi_{k_2})_{k_2,n} \tag{102}$$

that is, $e^{\mathcal{L}_n t}$ propagates chaos. The validity of the Boltzmann-Kac type equation (26) follows exactly as in [16]. □

4 Conclusions

The central aim of this work is the extension of the analysis in [4], in which a thermostat idealizes the interaction with a large reservoir of particles kept at constant temperature and chemical potential. While in [4] the reservoir and the system could not exchange particles, here the main interaction is the continuous exchange of particles between the two.

However, it is in this same work which we hoped to extend that we also find points of possible extension to our current work. In the case of the standard Kac model, approach to equilibrium in the sense of the GTW metric d_2 was shown in [18] while for a Kac system interacting with one or more Maxwellian thermostats it was shown in [8]. In the present situation though, it is not clear how to define an analogue of the GTW metric since the components f_N of a state $\mathbf{f}$ are not, in general, probability distributions on $\mathbb{R}^N$.

Furthermore, in [3] the authors show that, in a strong and uniform sense, the evolution of the Kac system with a Maxwellian thermostat can be thought of as an idealization of the interaction with a large heat reservoir, itself described as a Kac system. We think it is possible to replicate such an analysis in the present context and hope to come back to this issue in a forthcoming paper.

We based our proof of propagation of chaos on the work in [16]; therefore, as in [16], it is not quantitative nor uniform in time. Recently, a quantitative and uniform in time result was obtained for the Kac system with a Maxwellian thermostat [6]. It is unclear to us whether the methods in their work extend to the present model.

Finally, the assumption that the rates ρ and μ are independent of the number of particles is clearly unrealistic, allowing the possibility of an unbounded number of particles in the system. However, in the steady state (and in a chaotic state) the probability of having a number of particles in the system much larger then the average is extremely small, and so we do not consider this a serious problem. In any case, it would be interesting to investigate what happens if one assumes a maximum number of particles allowed inside the system.

Acknowledgements F.B. gratefully acknowledges National Science Foundation Grant DMS-1907643. F.B. thanks M. Loss and E. Carlen for many enlightening discussions and suggestions.

Data Availability Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

A Proofs of Technical Lemmas

A.1 Proof of Lemma 16

As already observed, we will search for the infimum of $(\mathbf{h}, -\mathcal{G}\mathbf{h})$ on $\mathbf{V}_m$ looking at states $\mathbf{h}$ close to $\mathcal{R}_m^+\mathbf{e}^0$. This is done using the representations (103) and (105) below. Since μ/ρ is large, (71) and (72) suggest that the dominant term in $(\mathbf{h}, \mathcal{G}\mathbf{h})$ for a state $\mathbf{h}$ close to $\mathcal{R}_m^+\mathbf{e}^0$ is the "diagonal term", that is the first term on the right hand side of (71) or (72). To prove Lemma 16 we thus need good bounds on the "off diagonal" terms. The proof in this section is thus loosely based on the proof of the Gershgorin circle theorem, see [10].

If $m = 2n + 1$, we can write any $\mathbf{h} \in \mathbf{V}_m$ as

$$\mathbf{h} = a\mathcal{R}_{2n+1}^+\mathbf{e}^0 + b\mathcal{R}_0^+\mathcal{R}_{2n+1}^+\mathbf{e}^0 + \mathbf{j} = a\mathcal{R}_{2n+1}^+\mathbf{e}^0 + \mathbf{k} \tag{103}$$

with $\mathbf{j} \perp \mathcal{R}_{2n+1}^+\mathbf{e}^0$ and $\mathbf{j} \perp \mathcal{R}_0^+\mathcal{R}_{2n+1}^+\mathbf{e}^0$.

From (72) we get $(\mathcal{R}_{2n+1}^+\mathbf{e}^0, \mathcal{K}\mathcal{R}_{2n+1}^+\mathbf{e}^0) = -\mu/\rho$ so that

$$\begin{aligned}(\mathbf{h}, -\widetilde{\mathcal{L}}\mathbf{h}) &= (\mathbf{h}, (-\mathcal{G} - \lambda\frac{\rho}{\mu}\mathcal{K})\mathbf{h}) \\ &= a^2(\lambda + \rho) + (\mathbf{k}, (-\mathcal{G} - \lambda\frac{\rho}{\mu}\mathcal{K})\mathbf{k}) + 2a(\mathcal{R}_{2n+1}^+\mathbf{e}^0, (-\mathcal{G} - \lambda\frac{\rho}{\mu}\mathcal{K})\mathbf{k})\end{aligned}$$

By construction $\mathbf{k}$ is in the span of the $\mathbf{e}_{\underline{\alpha}}$ with $\lambda(\underline{\alpha}) \ge 2$ so that $(\mathcal{R}_{2n+1}^+\mathbf{e}^0, \mathcal{G}\mathbf{k}) = 0$ and $(\mathbf{k}, -\mathcal{G}\mathbf{k}) \ge 2\rho\|\mathbf{k}\|_2 = 2\rho(b^2 + \|\mathbf{j}\|_2^2)$ while from (72) we get

$$(\mathcal{R}_{2n+1}^+\mathbf{e}^0, \mathcal{K}\mathbf{k}) = b(\mathcal{K}\mathcal{R}_{2n+1}^+\mathbf{e}^0, \mathcal{R}_{2n+1}^+\mathbf{e}^0) + (\mathcal{K}\mathcal{R}_{2n+1}^+\mathbf{e}^0, \mathbf{j}) = -\sqrt{\frac{\mu}{\rho}}b \tag{104}$$

This gives

$$\begin{aligned}(\mathbf{h}, -\widetilde{\mathcal{L}}\mathbf{h}) &\ge a^2(\lambda + \rho) + 2\rho(b^2 + \|\mathbf{j}\|_2^2) - 2\lambda|ab|\sqrt{\frac{\rho}{\mu}} \\ &\ge a^2\left(\rho + \lambda - \lambda\sqrt{\frac{\rho}{\mu}}\right) + b^2\left(2\rho - \lambda\sqrt{\frac{\rho}{\mu}}\right) + 2\rho\|\mathbf{j}\|_2^2 .\end{aligned}$$

Since $\|\mathbf{h}\|^2 = a^2 + b^2 + \|\mathbf{j}\|^2$ we get (73).

Similarly, every $\mathbf{h} \in \mathbf{V}_{2n}$ with $n \ge 2$ can be written as

$$\mathbf{h} = a\mathcal{R}_{2n}^+\mathbf{e}^0 + \sum_{k=0}^{n/2} b_k\mathcal{R}_{2k}^+\mathcal{R}_{2(n-k)}^+\mathbf{e}^0 + \mathbf{j} = a\mathcal{R}_{2n}^+\mathbf{e}^0 + \mathbf{k} \tag{105}$$

where $\mathbf{j} \perp \mathcal{R}_{2n}^+\mathbf{e}^0$ and $\mathbf{j} \perp \mathcal{R}_{2k}^+\mathcal{R}_{2(n-k)}^+\mathbf{e}^0$. Observe that

$$\left\|\sum_{k=0}^{n/2} b_k\mathcal{R}_{2k}^+\mathcal{R}_{2(nk)}^+\ be^0\right\|_2^2 = \sum_{k=0}^{n/2} \epsilon_{n,k}b_k^2$$

where, due to (60), $\epsilon_{n,k} = 2$ if $k = n - k$ and 1 otherwise. Analogously to (104), using (71), we get

$$(\mathcal{R}_{2n}^{+}\mathbf{e}^0, \mathcal{K}\mathcal{R}_{2n}^{+}\mathbf{e}^0) = \frac{\mu}{\rho}(2\tau_n - 1)$$

$$(\mathcal{R}_{2n}^{+}\mathbf{e}^0, \mathcal{K}\mathbf{k}) = 2\sqrt{\frac{\mu}{\rho}}\sum_{k=1}^{n/2} b_k\sigma_{n,k} + b_0\sqrt{\frac{\mu}{\rho}}(1 - 2\tau_n)\,.$$

Proceeding as before we obtain

$$\begin{aligned}
(\mathbf{h}, -\widetilde{\mathcal{L}}\mathbf{h}) &= (\mathbf{h}, (-\mathcal{G} - \lambda\frac{\rho}{\mu}\mathcal{K})\mathbf{h}) \\
&= a^2((1-2\tau_n)\lambda + \rho) + (\mathbf{k}, (-\mathcal{G} - \lambda\frac{\rho}{\mu}\mathcal{K})\mathbf{k}) + 2a(\mathcal{R}_{2n}^{+}\mathbf{e}^0, (-\mathcal{G} - \lambda\frac{\rho}{\mu}\mathcal{K})\mathbf{k}) \\
&\geq a^2((1-2\tau_n)\lambda + \rho) + 2\rho\left(\sum_{k=0}^{n/2}\epsilon_{n,k}b_k^2 + \|\mathbf{j}\|_2^2\right) \\
&\quad - 4\lambda\sqrt{\frac{\rho}{\mu}}\sum_{k=1}^{n/2}|ab_k|\sigma_{n,k} - 2\lambda|ab_0|\sqrt{\frac{\rho}{\mu}}(1-2\tau_n)\,.
\end{aligned}$$

which gives

$$\begin{aligned}
(\mathbf{h}, -\widetilde{\mathcal{L}}\mathbf{h}) &\geq a^2((1-2\tau_n)\lambda + \rho) + 2\rho\left(\sum_{k=0}^{n/2}\epsilon_{n,k}b_k^2 + \|\mathbf{j}\|^2\right) \\
&\quad - \lambda\sqrt{\frac{\rho}{\mu}}\left[2\sum_{k=1}^{n/2}a^2\sigma_{n,k}^2 + a^2(1-2\tau_n)^2 + 2\sum_{k=0}^{n/2}b_k^2\right] \\
&\geq \left(2\rho - 2\lambda\sqrt{\frac{\rho}{\mu}}\right)\left[\sum_{k=0}^{n/2}\epsilon_{n,k}b_k^2\right] + a^2\left((1-2\tau_n)\lambda + \rho - \lambda\sqrt{\frac{\rho}{\mu}}A_{2n}\right) + 2\rho\|\mathbf{j}\|_2^2
\end{aligned}$$

where

$$A_{2n} = (1-2\tau_n)^2 + 2\sum_{k=1}^{n/2}\sigma_{n,k}^2\,.$$

We thus need an upper bound on A_n. To this end, observe that

$$\log\tau_n = \log\prod_{i=1}^{n}\left(1 - \frac{1}{2i}\right) \leq -\sum_{i=1}^{n}\frac{1}{2i} \leq -\frac{1}{2}\log n \quad\Rightarrow\quad \tau_n \leq \frac{1}{\sqrt{n}}$$

while, from (70), we have

$$\sigma_{n,k}^2 \leq \tau_n\frac{1}{n}\frac{1}{\sqrt{\frac{k}{n}}\sqrt{1-\frac{k}{n}}}$$

so that

$$\sum_{k=1}^{n/2}\sigma_{n,k}^2 \leq \tau_n\int_0^{\frac{1}{2}}\frac{1}{\sqrt{x(1-x)}}dx = \frac{\pi}{2}\tau_n\,.$$

Finally we get $A_{2n} \leq 2$ which implies (74). □

A.2 Proof of Lemma 17

Proceeding as in (28) we can write

$$e^{\tilde{\lambda} K_N t} f_N = e^{-\tilde{\lambda}\binom{N}{2}t} f_N + t e^{-\tilde{\lambda}\binom{N}{2}t} \sum_{k=1}^{\infty} \frac{\tilde{\lambda}_n^k t^{k-1}}{k!} Q_N^k f_N$$

$$\frac{1}{t}\left(e^{\tilde{\lambda} K_N t} f_N - f_N\right) = e^{-\tilde{\lambda}\binom{N}{2}t} \tilde{\lambda} Q_N f_N + \frac{1}{t}\left(e^{-\tilde{\lambda}\binom{N}{2}t} - 1\right) f_N + t e^{-\tilde{\lambda}\binom{N}{2}t} \sum_{k=2}^{\infty} \frac{\tilde{\lambda}_n^k t^{k-2}}{k!} Q_N^k f_N$$

Since $R_N(\underline{v}_N, t) = \sum_{k=1}^{\infty} \frac{\tilde{\lambda}_n^k t^{k-1}}{k!}(Q_N^k f_N)(\underline{v}_N)$ is a sum of positive increasing terms and $\|R_N(t)\|_1 < \infty$, we see that $t e^{-\tilde{\lambda}\binom{N}{2}t} R_N(\underline{v}_N, t)$ converges to 0 as $t \to 0$ for almost every $\underline{v}_N$. A similar argument implies that $\left(e^{\tilde{\lambda} K_N t} f_N - f_N\right)/t$ converges almost everywhere to $\tilde{\lambda} K_N f_N$.

Using the Duhamel formula we can write

$$\begin{aligned}\left(e^{(\tilde{\lambda}\mathcal{K}+\rho(\mathcal{O}-\mathcal{N}))t}\mathbf{f}\right)_N &= e^{\left(\tilde{\lambda}K_N - \rho N\right)t} f_N \\ &\quad + \rho e^{\left(\tilde{\lambda}K_N - \rho N\right)t} \int_0^t e^{-\left(\tilde{\lambda}K_N - \rho N\right)s} \left(\mathcal{O} e^{\left(\tilde{\lambda}\mathcal{K}+\rho(\mathcal{O}-\mathcal{N})\right)s}\mathbf{f}\right)_N ds \\ &:= e^{\left(\tilde{\lambda}K_N - \rho N\right)t}\left(f_N + \overline{R}_N(t)\right). \end{aligned} \tag{106}$$

Since $\overline{R}_N(t, \underline{v}_N)$ is increasing in t and $\|\overline{R}_N(t)\|_{1,N} \to 0$ as $t \to 0^+$ we see that

$$\lim_{t \to 0^+} \left(e^{(\tilde{\lambda}\mathcal{K}+\rho(\mathcal{O}-\mathcal{N}))t}\mathbf{f}\right)_N (\underline{v}_N) = f_N(\underline{v}_N)$$

for almost every $\underline{v}_N$. Similarly using the Duhamel formula once more we get

$$\overline{R}_N(t) = \rho \int_0^t e^{-\left(\tilde{\lambda}K_N - \rho N\right)s} \left(\mathcal{O} e^{\left(\tilde{\lambda}\mathcal{K}-\rho\mathcal{N}\right)s}\mathbf{f}\right)_N ds + \overline{R}_{1,N}(t)$$

where

$$\overline{R}_{1,N}(t) = \rho^2 \int_0^t \int_0^s e^{-\left(\tilde{\lambda}K_N - \rho N\right)s} \left(\mathcal{O} e^{\left(\tilde{\lambda}\mathcal{K}-\rho\mathcal{N}\right)(s-s_1)} \mathcal{O} e^{\left(\tilde{\lambda}\mathcal{K}+\rho(\mathcal{O}-\mathcal{N})\right)s_1}\mathbf{f}\right)_N ds_1 ds$$

Reasoning as in (106) we get $\overline{R}_{1,N}(t, \underline{v}_N)/t \to 0$ as $t \to 0^+$ for almost every $\underline{v}_N$ while proceeding as in (39) we get

$$\lim_{t \to 0^+} \frac{1}{t} \int_0^t e^{-\left(\tilde{\lambda}K_N - \rho N\right)s} \left(\mathcal{O} e^{\left(\tilde{\lambda}\mathcal{K}-\rho\mathcal{N}\right)s}\mathbf{f}\right)_N ds = (\mathcal{O}\mathbf{f})_N \; .$$

Finally a similar argument using (40) concludes the proof. □

A.3 Proof of Lemma 18

Since $R_{i,j}$ is an average, we have $R_{i,j}h_N \log(R_{i,j}h_N) \leq R_{i,j}(h_n \log h_N)$, from which, calling $\overline{Q}_N = \binom{N}{2}^{-1} Q_N$, see (5), it follows that $\overline{Q}_N h_N \log(\overline{Q}_N h_N) \leq \overline{Q}_N(h_n \log h_N)$. Finally writing

$$e^{\tilde{\lambda} K_N} h_N = e^{-\tilde{\lambda}\binom{N}{2}t} \sum_{n=0}^{\infty} \binom{N}{2}^n \frac{\tilde{\lambda}^n t^n}{n!} \overline{Q}_N^n h_N$$

we get $e^{\tilde{\lambda} K_N} h_N \log(e^{\tilde{\lambda} K_N} h_N) \leq e^{\tilde{\lambda} K_N}(h_n \log h_N)$.

Proceeding as in Sect. 3.1 we can write

$$e^{\widetilde{\mathcal{L}}t}\mathbf{h} = e^{(\tilde{\lambda}\mathcal{K}-\rho\mathcal{N}-\mu\mathrm{Id})t}\mathbf{h} + \int_0^t e^{(\tilde{\lambda}\mathcal{K}-\rho\mathcal{N}-\mu\mathrm{Id})(t-s)}(\rho\mathcal{P}^+ + \mu\mathcal{P}^-)e^{\widetilde{\mathcal{L}}s}\mathbf{h}\, ds\,. \tag{107}$$

so that, writing $\mathbf{h}(t) = (h_0(t), h_1(v_1, t), \ldots)$ and using the notation introduced in the proof of Lemma 12, we get

$$\begin{aligned}
h_{N,t} &= e^{-(\rho N+\mu\mathrm{Id})t}\left(e^{\tilde{\lambda} K_N t} h_N(0)\right) \\
&\quad + \int_0^t e^{-(\rho N+\mu\mathrm{Id})(t-s)}\rho \sum_{i=1}^N e^{\tilde{\lambda} K_N (t-s)} P_{N,i}^+ h_{N-1}(s) ds \\
&\quad + \int_0^t e^{-(\rho N+\mu\mathrm{Id})(t-s)}\mu e^{\tilde{\lambda} K_N (t-s)} P_N^- h_{N+1}(s) ds
\end{aligned}$$

Observing that

$$e^{-(\rho N+\mu\mathrm{Id})t} + \int_0^t e^{-(\rho N+\mu\mathrm{Id})(t-s)}(\rho N + \mu) ds = 1$$

while

$$\begin{aligned}
P_{N,i}^+ h_{N-1} \log(P_{N,i}^+ h_{N-1}) &= P_{N,i}^+(h_{N-1} \log h_{N-1}) \\
P_N^- h_{N+1} \log(P_N^- h_{N+1}) &\leq P_N^-(h_{N+1} \log h_{N+1})
\end{aligned}$$

we get

$$\begin{aligned}
h_N(t) \log h_N(t) &\leq e^{(\tilde{\lambda} K_N - \rho N - \mu\mathrm{Id})t}(h_N(0) \log h_N(0)) \\
&\quad + \int_0^t e^{(\tilde{\lambda} K_N - \rho N - \mu\mathrm{Id})(t-s)}\rho \sum_{i=1}^N P_{N,i}^+(h_{N-1}(s) \log h_{N-1}(s)) ds \\
&\quad + \int_0^t e^{(\tilde{\lambda} K_N - \rho N - \mu\mathrm{Id})(t-s)}\mu P_N^-(h_{N+1}(s) \log h_{N+1}(s)) ds \\
&= \left(e^{(\tilde{\lambda}\mathcal{K}-\rho\mathcal{N}-\mu\mathrm{Id})t}\mathbf{h} \log \mathbf{h}\right)_N \\
&\quad + \left(\int_0^t e^{(\tilde{\lambda}\mathcal{K}-\rho\mathcal{N}-\mu\mathrm{Id})(t-s)}(\rho\mathcal{P}^+ + \mu\mathcal{P}^-)e^{\widetilde{\mathcal{L}}s}(\mathbf{h} \log \mathbf{h})\, ds\right)_N\,.
\end{aligned}$$

This, together with (107), completes the proof. □

A.4 Proof of Lemma 23

Given $h_N \in L^1(\mathbb{R}^N, \gamma_N)$ and a measurable set $A \in \mathbb{R}^N$

$$s(h_N) := \int_{\mathbb{R}^N} \gamma_N(\underline{v}_N) h_N(\underline{v}_N) \log h_N(\underline{v}_N) d\underline{v}_N \,.$$

and

$$\begin{aligned} m(A) &:= \int_A \gamma_N(\underline{v}_N) d\underline{v}_N \\ e(h_N, A) &:= \frac{1}{m(A)} \int_A h_N(\underline{v}_N)\gamma_N(\underline{v}_N) d\underline{v}_N \\ s(h_N, A) &:= \frac{1}{m(A)} \int_A h_N(\underline{v}_N) \log(h_N(\underline{v}_N))\gamma_N(\underline{v}_N) d\underline{v}_N \\ d(h_N, A) &:= s(h_N, A) - e(h_N, A) \log(e(h_N, A)) \,. \end{aligned}$$

Observe that m defines a probability measure on $\mathbb{R}^N$ while $d(h_N, A) \geq 0$ for every A.

Lemma 26 *Let $h_N > 0$ be such that $s(h_N) < \infty$. Then for every $\epsilon > 0$ there exists $\delta > 0$ such that, if A is a measurable set with $m(A) \leq \delta$ then $m(A)d(h_N, A) \leq \epsilon$.*

Proof Observe that

$$\begin{aligned} m(A)d(h_N, A) = &\int_A h_N(\underline{v}_N) \log(h_N(\underline{v}_N))\gamma_N(\underline{v}_N) d\underline{v}_N \\ &- \int_A h_N(\underline{v}_N)\gamma_N(\underline{v}_N) d\underline{v}_N \log\left(\int_A h_N(\underline{v}_N)\gamma_N(\underline{v}_N) d\underline{v}_N\right) \\ &+ \log(m(A)) \int_A h_N(\underline{v}_N)\gamma_N(\underline{v}_N) d\underline{v}_N \end{aligned} \tag{108}$$

The last term on the right hand side of (108) is negative while continuity of the Lebesgue integral implies that, given $\epsilon > 0$ we can find $\delta > 0$ such that, if A is a measurable set with $m(A) \leq \delta$ then the first and second terms in the right hand side of (108) are less then $\epsilon/2$. □

Given two partitions $\mathcal{B}_0$ and $\mathcal{B}_1$ of $\mathbb{R}$ we say that $\mathcal{B}_1$ *refines* $\mathcal{B}_0$ if every element of $\mathcal{B}_0$ can be written as a union of elements of $\mathcal{B}_1$. By convexity, if $\mathcal{B}_1$ refines $\mathcal{B}_0$ then $s(h_{N,\mathcal{B}_0^N}) \leq s(h_{N,\mathcal{B}_1^N})$. It is also easy to see that given two partitions $\mathcal{B}_0$ and $\mathcal{B}_1$ there always exists a partition $\mathcal{B}_2$ that refines both $\mathcal{B}_0$ and $\mathcal{B}_1$.

The following Lemma is the main result of this Appendix.

Lemma 27 *Let $h_N > 0$ be such that $s(h_N) < \infty$. Then for every ϵ there exists a finite partition $\mathcal{B}$ of $\mathbb{R}$ such that*

$$s(h_N) - s(h_{N,\mathcal{B}^N}) \leq \epsilon \,.$$

Proof Given a partition $\mathcal{B} = \{B_k\}_{k=1}^K$ of $\mathbb{R}$, we have

$$\begin{aligned} s(h_N) &= \sum_{\underline{k}\in\{1,\ldots,K\}^N} m(B_{\underline{k}}) s(h_N, B_{\underline{k}}) \\ s(h_{N,\mathcal{B}^N}) &= \sum_{\underline{k}\in\{1,\ldots,K\}^N} m(B_{\underline{k}}) e(h_N, B_{\underline{k}}) \log(e(h_N, B_{\underline{k}})) \end{aligned}$$

so that we need to find $\mathcal{B}$ such that

$$\sum_{\underline{k}\in\{1,\dots,K\}^N} m(B_{\underline{k}})d(h_N, B_{\underline{k}}) \le \epsilon\,.$$

To simplify notation, in what follows, we will write $d(A)$ for $d(h_N, A)$. Thanks to Lemma 26, given ϵ we can find δ such that for every A with $m(A) < \delta$ we have $m(A)d(A) < \epsilon/2$. Moreover there exists L such that calling $Q_L = (-L, L)^N$ we have $m(R) \le \delta/2$ for every $R \subset \mathbb{R}^N \setminus Q_L$.[1]

Let now $q_l(\underline{v}_N) = \{\underline{w}_N \,|\, |w_i - v_i| < l\}$, that is $q_l(\underline{v}_N)$ is the cube of side $2l$ centered at $\underline{v}_N$. By Lebesgue Differentiation Theorem, see e.g. Chap. 3 of [9], we get that $\lim_{l\to 0} d(q_l(\underline{v}_N)) = 0$ for m almost every $\underline{v}_N$. Thus, given ϵ_0 to be fixed later, there exists $\bar{l}$ such that $m(\{\underline{v}_N \,|\, d(q_{\bar{l}}(\underline{v}_N)) > \epsilon_0\}) \le \delta/2$. Let $Q_0 = \{\underline{v}_N \,|\, d(q_{\bar{l}}(\underline{v}_N)) \le \epsilon_0\} \cap Q_L$.

Let K be the smallest integer such that $l_0 = L/K < \bar{l}$ and consider the partition $\mathcal{A}$ of $[-L, L)$ formed by the sets $A_k = [kl_0, (k+1)l_0)$ with $-K \le k < K$. For every $A_{\underline{k}} \in \mathcal{A}^N$ let $C_{\underline{k}} = \emptyset$ if $A_{\underline{k}} \cap Q_0 = \emptyset$. Otherwise select a point $\underline{v}_N \in A_{\underline{k}} \cap Q_0$ and set $C_{\underline{k}} = q_{\bar{l}}(\underline{v}_N)$. Observe that, for every $\underline{k}$, $A_{\underline{k}} \cap Q_0 \subset C_{\underline{k}}$ so that $Q_0 \subset \bigcup_{\underline{k}} C_{\underline{k}} := Q_1$. This means that the $C_{\underline{k}}$ form a covering of Q_1 but not necessarily a partition. Let $\mathcal{D} = \{D_j\}_{j=1}^J$ be the minimal partition of Q_1 such that, for every j, $D_j \subset C_{\underline{k}}$ for some $\underline{k}$.[2] We claim that

$$\sum_{j=1}^J m(D_j)d(D_j) \le 5^N \epsilon_0 m(Q_1)\,. \tag{109}$$

To see this, let n_j be the number of $\underline{k}$ such that $D_j \subset C_{\underline{k}}$. By construction we have $n_j \ge 1$. On the other hand, for $\underline{v}_N \in A_{\underline{k}}$ and $\underline{w}_N \in A_{\underline{k}'}$, since $l_0 \ge \bar{l}/2$, we have $q_{\bar{l}}(\underline{v}_N) \cap q_{\bar{l}}(\underline{w}_N) = \emptyset$ if $\sum_{i=1}^N |k_i - k_i'| > 2$. This implies that $n_j \le 5^N$. Calling $J_{\underline{k}} = \{j \,|\, D_j \subset C_{\underline{k}}\}$, by convexity, we have

$$\sum_{j\in J_{\underline{k}}} m(D_j)d(D_j) \le m(C_{\underline{k}})d(C_{\underline{k}}) \le \epsilon_0 m(C_{\underline{k}})$$

because, by construction, the center of $C_{\underline{k}}$ is in Q_0. It follows that

$$\begin{aligned}\sum_{j=1}^J m(D_j)d(D_j) &\le \sum_{j=1}^J n_j m(D_j)d(D_j) = \sum_{\underline{k}} \sum_{j\in J_{\underline{k}}} m(D_j)d(D_j) \le \epsilon_0 \sum_{\underline{k}} m(C_{\underline{k}}) \\ &= \epsilon_0 \sum_{j=1}^J n_j m(D_j) \le \epsilon_0 5^N m(Q_1)\end{aligned}$$

where we used the bound on n_j and the fact that $\mathcal{D}$ is a partition.

We can now extend $\mathcal{D}$ to a partition $\widetilde{\mathcal{D}} = \{D_j\}_{j=1}^{\widetilde{J}}$ of $\mathbb{R}^N$ by adding finitely many rectangles. By construction, $m(\bigcup_{j=J+1}^{\widetilde{J}} D_j) = m(\mathbb{R}^N \setminus Q_1) \le \delta$ so that, choosing $\epsilon_0 = 5^{-N}\epsilon/2$, we get $s(h_N) - s(h_{N,\mathcal{D}}) \le \epsilon_0 5^N m(Q_1) + m(\mathbb{R}^N \setminus Q_1)d(\mathbb{R}^N \setminus Q_1) \le \epsilon$. Finally, since every $D_j \in \widetilde{\mathcal{D}}$ is a rectangle, we can find a finite partition $\mathcal{B}$ of $\mathbb{R}$ such that $\mathcal{B}^N$ refines $\widetilde{\mathcal{D}}$. This concludes the proof of Lemma 27. □

[1] Observe that for our purpose it is enough to work with a partition *mod* 0, that is a family of set B_k such that $m(\mathbb{R}\setminus \bigcup_k B_k) = 0$ and $m(B_k \cap B_{k'}) = 0$ for $k \neq k'$. For this reason, the boundaries of the rectangles defined in this proof are irrelevant.

[2] The partition $\mathcal{D}$ can be constructed by taking intersections of the $C_{\underline{k}}$ and their complements.

We are now ready to prove Lemma 23. Consider $\mathbf{h}$ such that $S(\mathbf{h}) < \infty$ and call

$$E_M(\mathbf{h}) = \sum_{N>M} a_N \int_{\mathbb{R}^N} \gamma(\underline{v}_N) h_N(\underline{v}_N) d\underline{v}_N$$
$$S_M(\mathbf{h}) = \sum_{N>M} a_N \int_{\mathbb{R}^N} \gamma(\underline{v}_N) h_N(\underline{v}_N) \log\left(h_N(\underline{v}_N)\right) d\underline{v}_N \, .$$

By convexity, for every M and every partition $\mathcal{B}$, we get

$$E_M(\mathbf{h}) \log(E_M(\mathbf{h})) \le S_M(\mathbf{h}_{\mathcal{B}}) \le S_M(\mathbf{h}) \, .$$

Since $E(\mathbf{h})$, $S(\mathbf{h}) < \infty$ for every ϵ there exists M such that $|E_M(\mathbf{h}) \log(E_M(\mathbf{h}))| \le \epsilon/4$ and $|S_M(\mathbf{h})| \le \epsilon/4$. This implies that, for every partition $\mathcal{B}$, we have $S_M(\mathbf{h}) - S_M(\mathbf{h}_{\mathcal{B}}) < \epsilon/2$. Moreover, from Lemma 27, for every $N \le M$ we can find a partition $\mathcal{B}_N$ of $\mathbb{R}$ such that

$$s(h_N) - s(h_{N,\mathcal{B}_N^N}) \le \frac{\epsilon}{2M} \, .$$

Finally let $\mathcal{B}$ be a partition of $\mathbb{R}$ that refines every $\mathcal{B}_N$ for $N \le M$. Since $s(h_{N,\mathcal{B}_N^N}) \le s(h_{N,\mathcal{B}^N}) \le s(h_N)$, we get

$$S(\mathbf{h}) = \sum_{N\le M} a_N s(h_N) + S_M(\mathbf{h}) \le \sum_{N\le M} a_N s(h_{N,\mathcal{B}^N}) + S_M(\mathbf{h}_{\mathcal{B}}) + \epsilon = S(\mathbf{h}_{\mathcal{B}}) + \epsilon$$

□

Remark 28 An alternative approach to coarse graining is as follows. Let I_s be the normalized characteristic function of the segment $(-s, s)$ and let $(\mathbf{h}_s)_N = \int I_s^{\otimes N}(\underline{v}_N - \underline{w}_N) h_N(\underline{w}_N) d\underline{w}_N$. Clearly $(\mathbf{h}_s)_N$ is continuous for every N. Moreover we have $S(\mathbf{h}_s) \le S(\mathbf{h})$ and $\Psi(\mathbf{h}_s) \le \Psi(\mathbf{h})$, see the argument around (76). Finally, $S(\mathbf{h}_s) \to_{s\to 0} S(\mathbf{h})$. Thus, reasoning like in (78), we can restrict our attention to continuous states $\mathbf{h}$. In this case, the analogue of Lemma 27 is simpler to prove. Indeed if h_N is continuous, $\lim_{l\to 0} d(q_l(\underline{v}_N), h_N) = 0$ for every $\underline{v}_N$ and, since Q_L is compact, we can find $\bar{l}$ such that $d(q_{\bar{l}}(\underline{v}_N), h_N) \le \epsilon$ for $\underline{v}_N \in Q_L$. In such a situation, the regular partition $\mathcal{A}$ built before (109) already provides the solution.

A.5 Derivation of (98) and (102)

We start with (98). Expanding the terms in (33) the form $(e^{(\tilde{\lambda}\mathcal{K}-\rho\mathcal{N})t}\mathbf{f}_n, \phi_k)_{k,n}$ using (97) recursively starting from the most external one gives

$$\lim_{n\to\infty} \left(e^{(\tilde{\lambda}_n\mathcal{K}+\rho(\mathcal{O}-\mathcal{N}))t}\mathbf{f}_n, \phi_k\right)_{k,n} \tag{110}$$
$$= \lim_{n\to\infty} \sum_{q\ge 0} \sum_{p_0,p_1,\ldots,p_q\ge 0} \rho^q \lambda^{|p|} \left(\frac{\rho}{\mu_n}\right)^{k+|p|} \sum_{N\ge k+|p|} \frac{N!}{(N-k-|p|)!} \frac{(N+q)!}{N!} \tag{111}$$
$$\cdot \int_{0\le t_q\le\cdots t_1\le t} \prod_{i=0}^{q} \left(e^{-\rho(N+q-i)(t_i-t_{i+1})} \frac{(t_i-t_{i+1})^{p_i}}{p_i!}\right) dt_1\cdots dt_q \tag{112}$$
$$\cdot \int f_{N+q}(\underline{v}_{N+q})(G_k^{*|p|}\phi_k)(\underline{v}_{k+|p|}) d\underline{v}_{N+q} \tag{113}$$

where $t_{q+1} = 0$, $t_0 = t$ and $|p| = \sum_{i=0}^{q} p_i$. Call now

$$b_{q,P} = \sum_{\substack{p_0,p_1,\dots,p_q\geq 0\\ |p|=P}} \int_{0\leq t_q\leq\cdots t_1\leq t} \prod_{i=0}^{q} \left(e^{-\rho(N+q-i)(t_i-t_{i+1})} \frac{(t_i - t_{i+1})^{p_i}}{p_i!} \right) dt_1 \cdots dt_q \,.$$

We first sum over the p_i using that

$$\sum_{\substack{p_0,p_1,\dots,p_q\geq 0\\ |p|=P}} \prod_{i=0}^{q} \frac{(t_i - t_{i+1})^{p_i}}{p_i!} = \frac{t^p}{p!} \,,$$

then we integrate over the t_i using (35) and we get

$$b_{q,P} = \frac{t^p}{p!}\frac{1}{q!} e^{-\rho N t}(1 - e^{-\rho t})^q$$

Inserting in (110) gives

$$\lim_{n\to\infty} \left(e^{(\tilde{\lambda}_n \mathcal{K} + \rho(\mathcal{O}-\mathcal{N}))t} \mathbf{f}_n, \phi_k \right)_{k,n} = \lim_{n\to\infty} \sum_{q\geq 0}\sum_{p\geq 0} \frac{t^p \lambda^p}{p!} \left(\frac{\rho}{\mu_n}\right)^{k+p}$$
$$\cdot \sum_{N\geq k+p} \frac{(N+q)!}{(N-k-p)!\,q!} e^{-\rho N t}(1-e^{-\rho t})^q \int f_{N+q}(\underline{v}_{N+q})(G_k^{*|p|}\phi_k)(\underline{v}_{k+|p|})d\underline{v}_{N+q} \,.$$

Finally we write

$$\frac{(N+q)!}{(N-k-p)!\,q!} = \frac{(N+q)!}{(N+q-k-p)!} \binom{N+q-k-p}{q}$$

so that, setting $M = N + q$ and summing over q, we get

$$\lim_{n\to\infty} \left(e^{(\tilde{\lambda}_n \mathcal{K} + \rho(\mathcal{O}-\mathcal{N}))t} \mathbf{f}_n, \phi_k \right)_{k,n} = \lim_{n\to\infty} \sum_{p\geq 0} \frac{t^p \lambda^p}{p!} e^{-\rho(k+p)t}$$
$$\left(\frac{\rho}{\mu_n}\right)^{k+p} \sum_{M\geq k+p} \frac{M!}{(M-k-p)!} \int f_{N+q}(\underline{v}_M)(G_k^{*|p|}\phi_k)(\underline{v}_M)d\underline{v}_{N+q} \,.$$

and the derivation of (98) is complete.

Turning to (102) we set

$$\mathcal{A}_q(\phi_k) = \sum_{p_0,p_1,\dots,p_q\geq 0} \int_{0\leq t_q\leq\cdots\leq t_1\leq t} \prod_{i=0}^{q} e^{-\rho(t_i-t_{i+1})(|p|_i - i)} \frac{(t_i - t_{i+1})^{p_i}}{p_i!} dt_1 \cdots dt_q$$
$$\cdot G^{*p_q}_{k+|p|_q-q} I_{k+|p|_q-q+1} \cdots G^{*p_1}_{k+p_0-1} I_{k+p_0} G_k^{*p_0} \phi_k$$

For the rest of this section we will neglect the number of variables subscript in order to make expressions more readable. To understand the structure of $\mathcal{A}_q(\phi \otimes \psi)$, first look at $\mathcal{A}_2(\phi \otimes \psi)$. Combining (94) and (99) we can write

$$\sum_{p_0,p_1,p_2} \prod_{i=0}^{2} \frac{(e^{-(t_i-t_{i+1})}(t_i - t_{i+1}))^{p_i}}{p_i!} G^{*p_2} I G^{*p_1} I G^{*p_0}(\phi \otimes \psi)$$

$$= \sum_{p_0^1,p_1^1,p_2^1} \sum_{p_0^2,p_1^2,p_2^2} \prod_{i=0}^{2} \frac{(e^{-(t_i-t_{i+1})}(t_i - t_{i+1}))^{p_i^1}}{p_i^1!} \prod_{i=0}^{2} \frac{(e^{-(t_i-t_{i+1})}(t_i - t_{i+1}))^{p_i^2}}{p_i^2!}$$

$$\Big(G^{*p_2^1} I G^{*p_1^1} I G^{*p_0^1}\phi \otimes G^{*p_2^2} G^{*p_1^2} G^{*p_0^2}\psi \tag{114a}$$

$$+ G^{*p_2^1} I G^{*p_1^1} G^{*p_0^1}\phi \otimes G^{*p_2^2} G^{*p_1^2} I G^{*p_0^2}\psi \tag{114b}$$

$$+ G^{*p_2^1} G^{*p_1^1} I G^{*p_0^1}\phi \otimes G^{*p_2^2} I G^{*p_1^2} G^{*p_0^2}\psi \tag{114c}$$

$$+G^{*p_2^1} G^{*p_1^1} G^{*p_0^1}\phi \otimes G^{*p_2^2} I G^{*p_1^2} I G^{*p_0^2}\psi\Big)\,. \tag{114d}$$

To simplify (114a), we can use that

$$\sum_{p_0^2,p_1^2,p_2^2} \prod_{i=0}^{2} \frac{(e^{-(t_i-t_{i+1})}(t_i - t_{i+1}))^{p_i^2}}{p_i^2!} (\mathbf{f}_n, G^{*p_2^2} G^{*p_1^2} G^{*p_0^2}\psi)_n = \sum_p \frac{t^p}{p!} e^{-t} (\mathbf{f}_n, G^{*p}\psi)_n$$

and similarly for (114d). On the other hand, for (114b) we have

$$\sum_{p_0^1,p_1^1,p_2^1} \sum_{p_0^2,p_1^2,p_2^2} \prod_{i=0}^{2} \frac{(e^{-(t_i-t_{i+1})}(t_i - t_{i+1}))^{p_i^1}}{p_i^1!} \prod_{i=0}^{2} \frac{(e^{-(t_i-t_{i+1})}(t_i - t_{i+1}))^{p_i^2}}{p_i^2!}$$
$$(\mathbf{f}_n, G^{*p_2^1} I G^{*p_1^1} G^{*p_0^1}\phi)_n (\mathbf{f}_n, G^{*p_2^2} G^{*p_1^2} I G^{*p_0^2}\psi)_n$$
$$= \sum_{p_0^1,p_1^1} \sum_{p_0^2,p_1^2} \frac{(e^{-(t-t_1)}(t-t_1))^{p_0^1} (e^{-t_1}t_1)^{p_1^1}}{p_0^1! p_1^1!} \frac{(e^{-(t-t_2)}(t-t_2))^{p_0^2} (e^{-t_2}t_2)^{p_1^2}}{p_0^2! p_1^2!}$$
$$(\mathbf{f}_n, G^{*p_1^1} I G^{*p_0^1}\phi)_n (\mathbf{f}_n, G^{*p_1^2} I G^{*p_0^2}\psi)_n\,.$$

while (114c) gives a similar expression but for the roles of t_1 and t_2 that are inverted. Combining these expressions we get

$$\mathcal{A}_2(\phi \otimes \psi) = \mathcal{A}_0(\phi)\mathcal{A}_2(\psi) + \mathcal{A}_1(\phi)\mathcal{A}_1(\psi) + \mathcal{A}_2(\phi)\mathcal{A}_0(\psi)$$

For the general case we can write

$$\prod_{i=0}^{q} e^{-\rho(t_i-t_{i+1})(k+|p|_i-i)} \frac{(t_i - t_{i+1})^{p_i}}{p_i!} G^{*p_{q+1}} I \cdots G^{*p_1} I G^{*p_0}(\phi \otimes \psi)$$
$$= \sum_{p_i^1+p_i^2=p_i} \sum_{\sigma_1,\ldots\sigma_q\in\{0,1\}}$$
$$\prod_{i=0}^{q} e^{-\rho(t_i-t_{i+1})(k_1+|p^1|_i-i)} \frac{(t_i - t_{i+1})^{p_i^1}}{p_i^1!} G^{*p_{q+1}^1} I^{\sigma_q} \cdots G^{*p_1^1} I^{\sigma_1} G^{*p_0^1}\phi$$
$$\otimes \prod_{i=0}^{q} e^{-\rho(t_i-t_{i+1})(k_2+|p^2|_i-i)} \frac{(t_i - t_{i+1})^{p_i^2}}{p_i^2!} G^{*p_{q+1}^2} I^{1-\sigma_q} \cdots G^{*p_1^2} I^{1-\sigma_1} G^{*p_0^2}\psi\,.$$

that, after resummation, gives

$$A_q(\phi \otimes \psi) = \sum_{q_1+q_2=q} \sum_{p_0^1,p_1^1,\ldots,p_{q^1+1}^1\geq 0} \sum_{p_0^2,p_1^2,\ldots,p_{q^2+1}^2\geq 0} \lambda^{|p^1|+|p^2|}$$

$$
\begin{aligned}
&\sum_{\pi}^{*}\int_{0\le t_{\pi(q_1+q_2)}\le\cdots\le t_{\pi(1)}\le t} dt_{1,1}\cdots dt_{1,q_1}dt_{2,1}\cdots dt_{2,q_2}\\
&\prod_{i=0}^{q_1} e^{-\rho(t_{1,i}-t_{1,i+1})(k_1+|p^1|_i-i)}\frac{(t_{1,i}-t_{1,i+1})^{p_i^1}}{p_i^1!}\quad G^{*p^1_{q_1+1}}I\cdots G^{*p^1_1}IG_{k_1}^{*p^1_0}\phi\\
&\otimes\prod_{j=0}^{q_2} e^{-\rho(t_{2,j}-t_{2,j+1})(k_2+|p^2|_j-j)}\frac{(t_{2,j}-t_{2,j+1})^{p_j^2}}{p_j^2!}G^{*p^2_{q_2+1}}I\cdots G^{*p^2_1}IG^{*p^2_0}\psi
\end{aligned}
$$

where $\sum_{\pi}^{*}$ is the sum over all one-to-one functions π from $\{1,\ldots,q_1+q_2\}$ to the set $\{(1,1),\ldots,(1,q_1),(2,1),\ldots,(2,q_2)\}$ such that if, for $i>j$ and $\sigma\in\{1,2\}$, we have $\pi(i)=(\sigma,q)$ and $\pi(j)=(\sigma,q')$ then $q>q'$. Observing that

$$
\begin{aligned}
&\sum_{\pi}^{*}\int_{0\le t_{\pi(1)}\le\cdots\le t_{\pi(q_1+q_2)}\le t} dt_{1,1}\cdots dt_{1,q_1}dt_{2,1}\cdots dt_{2,q_2}\\
&\quad=\int_{0\le t_{1,1}\le\cdots\le t_{1,q_1}\le t} dt_{1,q_1}\cdots dt_{1,1}\int_{0\le t_{2,q_2}\le\cdots\le t_{2,1}\le t} dt_{2,1}\cdots dt_{2,q_2}
\end{aligned}
$$

we get

$$
\mathcal{A}_q(\phi\otimes\psi)=\sum_{q_1+q_2=q}\mathcal{A}_{q_1}(\phi)\mathcal{A}_{q_2}(\psi)\,.
$$

Propagation of chaos now follows easily.

References

1. Benfatto, G., Gallavotti, G.: Perturbation theory of the fermi surface in a quantum liquid. A general quasiparticle formalism and one-dimensional systems. J. Stat. Phys. **59**, 541–664 (1990)
2. Bonetto, F., Geisinger, A., Loss, M., Ried, T.: Entropy decay for the Kac evolution. Commun. Math. Phys. **363**(3), 847–875 (2018). https://doi.org/10.1007/s00220-018-3263-0
3. Bonetto, F., Loss, M., Tossounian, H., Vaidyanathan, R.: Uniform approximation of a Maxwellian thermostat by finite reservoirs. Commun. Math. Phys. **351**(1), 311–339 (2017). https://doi.org/10.1007/s00220-016-2803-8
4. Bonetto, F., Loss, M., Vaidyanathan, R.: The Kac model coupled to a thermostat. J. Stat. Phys. **156**(4), 647–667 (2014). https://doi.org/10.1007/s10955-014-0999-6
5. Carlen, E., Carvalho, M.C., Loss, M.: Many-body aspects of approach to equilibrium. In: Journées "Équations aux Dérivées Partielles" (La Chapelle sur Erdre, 2000), pp. Exp. No. XI, 12. Univ. Nantes, Nantes (2000)
6. Cortez, R., Tossounian, H.: Uniform propagation of chaos for the thermostated Kac model. J. Stat. Phys. **183**, 28 (2021)
7. Einav, A.: On Villani's conjecture concerning entropy production for the Kac master equation. Kinet. Relat. Models **4**(2), 479–497 (2011). https://doi.org/10.3934/krm.2011.4.479
8. Evans, J.: Non-equilibrium steady states in Kac's model coupled to a thermostat. J. Stat. Phys. **164**, 1103–1121 (2016)
9. Folland, G.B.: Real Analysis: Modern Techniques and Their Applications. Wiley, New York (1999)
10. Golub, G.H., Van Loan, C.F.: Matrix Computations. Johns Hopkins University Press, Baltimore (1996)
11. Grad, H.: On the kinetic theory of rarefied gases. Commun. Pure Appl. Math. **2**, 331–407 (1949)
12. Gross, L.: Logarithmic Sobolev inequalities. Am. J. Math. **97**(4), 1061–1083 (1975)
13. Janvresse, E.: Spectral gap for Kac's model of Boltzmann equation. Ann. Probab. **29**(1), 288–304 (2001). https://doi.org/10.1214/aop/1008956330

14. 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 and Los Angeles (1956)
15. Kac, M.: Probability and related topics in physical sciences, with special lectures by G. E. Uhlenbeck, A.R. Hibbs, and B. van der Pol. Lectures in Applied Mathematics. In: Proceedings of the Summer Seminar, Boulder, Colo., vol. 1957. Interscience Publishers, London (1959)
16. McKean, H.P., Jr.: Speed of approach to equilibrium for Kac's caricature of a Maxwellian gas. Arch. Ration. Mech. Anal. **21**, 343–367 (1966)
17. Schweber, S.: An Introduction to Relativistic Quantum Field Theory. Dover Publications, Mineola (2011)
18. Tossounian, H.: Equilibration in the Kac model using the GTW metric d2. J. Stat. Phys. **169**, 168–186 (2017)
19. Villani, C.: Cercignani's conjecture is sometimes true and always almost true. Commun. Math. Phys. **234**(3), 455–490 (2003). https://doi.org/10.1007/s00220-002-0777-1