Communications in
**Mathematical
Physics**

CrossMark

Superdiffusion in the Periodic Lorentz Gas

Jens Marklof[1], Bálint Tóth[1,2,3]

[1] School of Mathematics, University of Bristol, Bristol BS8 1TW, UK. E-mail: j.marklof@bristol.ac.uk
[2] MTA-BME Stochastics Research Group, Budapest, Hungary.
[3] Rényi Institute, Budapest, Hungary. E-mail: balint.toth@bristol.ac.uk; balint@math.bme.hu

Abstract: We prove a superdiffusive central limit theorem for the displacement of a test particle in the periodic Lorentz gas in the limit of large times t and low scatterer densities (Boltzmann–Grad limit). The normalization factor is $\sqrt{t \log t}$, where t is measured in units of the mean collision time. This result holds in any dimension and for a general class of finite-range scattering potentials. We also establish the corresponding invariance principle, i.e., the weak convergence of the particle dynamics to Brownian motion.

1. Introduction

The periodic Lorentz gas is one of the iconic models of "chaotic" diffusion in deterministic systems. It describes the dynamics of a test-particle in an infinite periodic array of spherically symmetric scatterers. The main results characterizing the diffusive nature of the periodic Lorentz gas have to date been mainly restricted to the two-dimensional setting and hard-sphere scatterers. The first seminal result on this subject was the proof of a central limit theorem for the displacement of the test particle at large times t for the finite-horizon Lorentz gas by Bunimovich and Sinai [9]. For more general invariance principles see Melbourne and Nicol [23] and references therein. In the case of the infinite-horizon Lorentz gas, Bleher [6] pointed out that the mean-square displacement grows like $t \log t$ when $t \to \infty$, as opposed to a linear growth in the finite-horizon case. The superdiffusive central limit theorem suggested in [6] was first proved by Szász and Varjú [29] for the discrete-time billiard map. Dolgopyat and Chernov [15]

provided an alternative proof, and established the central limit theorem and invariance principle for the billiard flow. Analogous results hold for the stadium billiard (Bálint and Gouëzel [2]) and billiards with cusps (Bálint et al. [1]). The difficulty in extending the above findings to dimensions greater than two lies in the possibly exponential growth of the complexity of singularities (Bálint and Tóth [3,4], Chernov [13]) and, in the case of infinite horizon, the subtle geometry of channels (Dettmann [14], Nándori et al. [24]).

In the present paper we prove unconditional superdiffusive central limit theorems and invariance principles for the periodic Lorentz gas in any dimension $d \geq 2$, valid in the limit of low scatterer density (Boltzmann–Grad limit) and for a general class of finite-range scattering potentials. That is, instead of fixing the radius r of each scatterer and considering the long time limit as in the above cited papers, we consider here the limit $r \to 0$ and then the limit of long times, where time is measured in units of the mean collision time. It is an interesting open problem to consider the two limits $r \to 0$, $t \to \infty$ jointly.

The precise setting of our study is as follows. Let $\mathcal{L} \subset \mathbb{R}^d$ be a fixed Euclidean lattice of covolume one (such as the cubic lattice $\mathcal{L} = \mathbb{Z}^d$), and define the scaled lattice $\mathcal{L}_r := r^{(d-1)/d}\mathcal{L}$. At each point in $\mathcal{L}_r$ we center a sphere of radius r. We consider a test particle that moves along straight lines with unit speed until it hits a sphere, where it is scattered elastically. The above scaling of scattering radius vs. lattice spacing ensures that the mean free path length (i.e., the average distance between consecutive collisions) has the limit $\overline{\xi} = 1/v_{d-1}$ as $r \to 0$, where $v_{d-1} = \pi^{\frac{d-1}{2}}/\Gamma(\frac{d+1}{2})$ denotes the volume of the unit ball in $\mathbb{R}^{d-1}$.

In the case of the classic Lorentz gas the scattering mechanism is given by specular reflection, but as in [21] we will here also allow more general spherically symmetric scattering maps. The precise conditions will be stated in Sect. 2.

The position of our test particle at time t is denoted by

$$\boldsymbol{x}_t = \boldsymbol{x}_t(\boldsymbol{x}_0, \boldsymbol{v}_0) \in \mathcal{K}_r := \mathbb{R}^d \setminus (\mathcal{L}_r + r\mathcal{B}_1^d), \tag{1.1}$$

where $\boldsymbol{x}_0$ and $\boldsymbol{v}_0$ are the position and velocity at time $t = 0$, and $\mathcal{B}_1^d$ is the open unit ball in $\mathbb{R}^d$ centered at the origin. We use the convention that for any boundary point $\boldsymbol{x}_0 \in \partial\mathcal{K}_r$ we choose the *outgoing* velocity $\boldsymbol{v}_0$, i.e., the velocity *after* the scattering. The corresponding phase space is denoted by $\mathrm{T}^1(\mathcal{K}_r)$. For notational reasons it is convenient to extend the dynamics to $\mathrm{T}^1(\mathbb{R}^d) := \mathbb{R}^d \times \mathrm{S}_1^{d-1}$ by setting $\boldsymbol{x}_t = \boldsymbol{x}_0$ for all initial conditions $\boldsymbol{x}_0 \notin \mathcal{K}_r$.

We consider the time evolution of a test particle with random initial data $(\boldsymbol{x}_0, \boldsymbol{v}_0) \in \mathrm{T}^1(\mathbb{R}^d)$, distributed according to a given Borel probability measure Λ on $\mathrm{T}^1(\mathbb{R}^d)$. The following superdiffusive central limit theorem, valid for small scattering radii and large times, asserts that the normalized particle displacement at time t, and measured in units of $\sqrt{t \log t}$, converges weakly to a Gaussian distribution.

Theorem 1.1. *Let $d \geq 2$ and fix a Euclidean lattice $\mathcal{L} \subset \mathbb{R}^d$ of covolume one. Assume $(\boldsymbol{x}_0, \boldsymbol{v}_0)$ is distributed according to an absolutely continuous Borel probability measure Λ on $\mathrm{T}^1(\mathbb{R}^d)$. Then, taking first $r \to 0$ and then $t \to \infty$, we have*

$$\frac{\boldsymbol{x}_t - \boldsymbol{x}_0}{\Sigma_d \sqrt{t \log t}} \Rightarrow \mathcal{N}(0, I_d), \tag{1.2}$$

where $\mathcal{N}(0, I_d)$ is a centered normal random variable in $\mathbb{R}^d$ with identity covariance matrix, and

$$\Sigma_d^2 := \frac{2^{1-d} v_{d-1}}{d^2(d+1)\zeta(d)}. \tag{1.3}$$

Here $\zeta(d) := \sum_{n=1}^{\infty} n^{-d}$ denotes the Riemann zeta function. Recall that the weak convergence (1.2) holds if and only if

$$\lim_{t\to\infty} \lim_{r\to 0} \mathbf{E} \, f\left(\frac{x_t - x_0}{\Sigma_d \sqrt{t \log t}}\right) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} f(x) \, e^{-\frac{1}{2}\|x\|^2} dx \tag{1.4}$$

for any bounded continuous $f : \mathbb{R}^d \to \mathbb{R}$.

Theorem 1.1 will follow from its discrete-time analogue, Theorem 1.2. Let us denote by $q_n = q_n(q_0, v_0) \in \partial \mathcal{K}_r$ $(n = 1, 2, 3, \ldots)$ the location where the test particle with initial condition (q_0, v_0) leaves the nth scatterer. It is natural in this setting to assume $q_0 \in \partial \mathcal{K}_r$. By the translational invariance of the lattice, we may in fact assume without loss of generality $q_0 \in r S_1^{d-1}$. For given exit velocity v_0, we write

$$q_0 = r(s_0 + v_0 \sqrt{1 - \|s_0\|^2}) \tag{1.5}$$

and stipulate in the following that the random variable s_0 is uniformly distributed in the unit disc orthogonal to v_0. The uniform distribution is the natural invariant measure for the discrete time dynamics.

Theorem 1.2. *Let $d \geq 2$ and $\mathcal{L}$ as above. Assume v_0 is distributed according to an absolutely continuous Borel probability measure λ on S_1^{d-1}. Then, taking first $r \to 0$ and then $n \to \infty$, we have*

$$\frac{q_n - q_0}{\sigma_d \sqrt{n \log n}} \Rightarrow \mathcal{N}(0, I_d), \tag{1.6}$$

with

$$\sigma_d^2 := \frac{2^{1-d}}{d^2(d+1)\zeta(d)} = \overline{\xi} \, \Sigma_d^2. \tag{1.7}$$

The above results generalise to functional central limit theorems, also known as invariance principles. Denote by $C_0([0, 1])$ the space of curves $[0, 1] \to \mathbb{R}^d$ starting at the origin. We fix a metric on $C_0([0, 1])$ by defining the distance between two curves X_1 and X_2 by $\sup_{t\in[0,1]} \|X_1(t) - X_2(t)\|$. The topology generated by open balls in this metric is called the *uniform topology*. A sequence $(X_n)_n$ of random curves in $C_0([0, 1])$ *converges weakly* to X ($X_n \Rightarrow X$), if for any bounded continuous $f : C_0([0, 1]) \to \mathbb{R}$ we have $\lim_n \mathbf{E} \, f(X_n) = \mathbf{E} \, f(X)$.

The following theorem, which is the main result of this paper, states that for the same random initial data as in Theorem 1.1, the random curves

$$[0, 1] \to \mathbb{R}^d, \quad t \mapsto X_{T,r}(t) := \frac{x_{tT} - x_0}{\Sigma_d \sqrt{T \log T}}, \tag{1.8}$$

converge weakly to the standard Brownian motion $t \mapsto W(t)$ in $\mathbb{R}^d$ with unit covariance matrix I_d.

Theorem 1.3. *Let $d \geq 2$ and fix a Euclidean lattice $\mathcal{L} \subset \mathbb{R}^d$ of covolume one. Assume $(\boldsymbol{x}_0, \boldsymbol{v}_0)$ is distributed according to an absolutely continuous Borel probability measure Λ on $\mathrm{T}^1(\mathbb{R}^d)$. Then, taking first $r \to 0$ and then $T \to \infty$, we have*

$$X_{T,r} \Rightarrow W. \tag{1.9}$$

As in the case of Theorem 1.1, we derive Theorem 1.3 as a corollary of its discrete-time analogue, Theorem 1.4. By linearly interpolating between the position variables $\boldsymbol{q}_0, \boldsymbol{q}_1, \ldots, \boldsymbol{q}_n$, we obtain the piecewise linear curve

$$[0, 1] \to \mathbb{R}^d, \quad t \mapsto \boldsymbol{q}_n(t) := \boldsymbol{q}_{\lfloor nt \rfloor} + \{nt\} \left(\boldsymbol{q}_{\lfloor nt \rfloor + 1} - \boldsymbol{q}_{\lfloor nt \rfloor} \right), \tag{1.10}$$

where $\{x\} := x - \lfloor x \rfloor$ denotes the fractional part of x. We rescale the curve by setting

$$\boldsymbol{Y}_{n,r}(t) := \frac{\boldsymbol{q}_n(t) - \boldsymbol{q}_0}{\sigma_d \sqrt{n \log n}}. \tag{1.11}$$

We then have the following generalization of Theorem 1.2.

Theorem 1.4. *Let $d \geq 2$ and $\mathcal{L}$ a Euclidean lattice of covolume one. Assume $\boldsymbol{v}_0$ is distributed according to an absolutely continuous Borel probability measure λ on S_1^{d-1}. Then, taking first $r \to 0$ and then $n \to \infty$, we have*

$$Y_{n,r} \Rightarrow W. \tag{1.12}$$

The starting point of our analysis is the paper [21], which proves that, for every fixed $t > 0$, the limit $r \to 0$ in (1.2) (resp. (1.6)) exists and is given by a continuous-time (resp. discrete-time) Markov process. The main objective of the present study is therefore to prove a superdiffusive central limit theorem, as well as an invariance principle, for each of these Markov processes. The central limit theorem is stated as Theorem 3.2 in Sect. 3 after a brief survey of the relevant results from [21]. The subsequent sections of the paper are devoted to the proof of Theorem 3.2. The invariance principles stated in Theorems 1.3 and 1.4 follow from the results in Sects. 12–14.

2. The Scattering Map

We now specify the conditions on the scattering map that are assumed in Theorems 1.1–1.4. These are the same as in [21], with the additional simplifying assumption that the scattering map preserves angular momentum, cf. [21, Remark 2.3]. We describe the scattering map in units of r, i.e., the scatterer is represented as the open unit ball $\mathcal{B}_1^d$. Set

$$\mathcal{S} := \{(\boldsymbol{v}, \boldsymbol{b}) \in \mathrm{S}_1^{d-1} \times \mathcal{B}_1^d \mid \boldsymbol{v} \cdot \boldsymbol{b} = 0\}, \tag{2.1}$$

and consider the scattering map

$$\Theta : \mathcal{S} \to \mathcal{S}, \quad (\boldsymbol{v}_-, \boldsymbol{b}) \mapsto (\boldsymbol{v}_+, \boldsymbol{s}). \tag{2.2}$$

The incoming data is denoted by $(\boldsymbol{v}_-, \boldsymbol{b}) \in \mathcal{S}$, where $\boldsymbol{v}_-$ is the velocity of the particle before the collision and $\boldsymbol{b}$ the impact parameter, i.e., the point of impact on S_1^{d-1} projected onto the plane $\{\boldsymbol{b} \in \mathbb{R}^d \mid \boldsymbol{v}_- \cdot \boldsymbol{b} = 0\}$. The outgoing data is analogously defined as $(\boldsymbol{v}_+, \boldsymbol{s}) \in \mathcal{S}$, where $\boldsymbol{v}_+$ is the velocity of the particle after the collision and $\boldsymbol{s}$ the exit parameter, cf. Figure 1. Since we assume the scattering map is spherically symmetric,
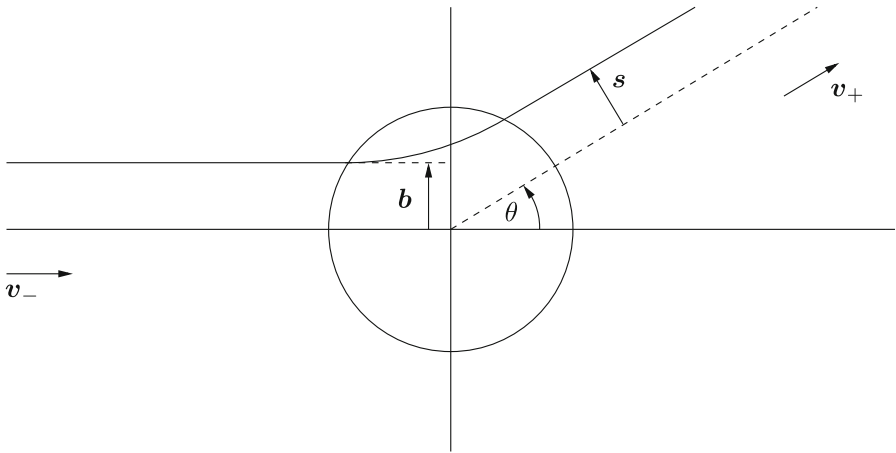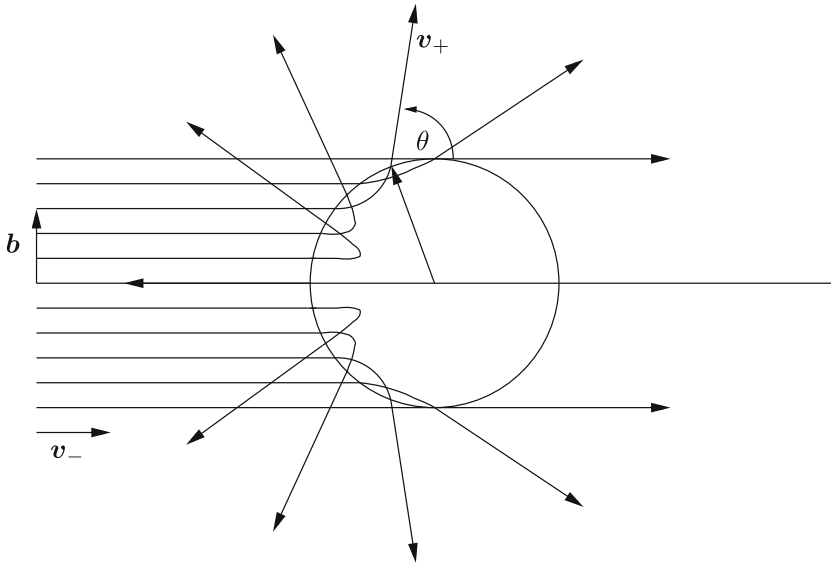
Fig. 1. The scattering map

Fig. 2. Illustration of a scattering map satisfying Hypothesis (A)

it is sufficient to define Θ for $(\boldsymbol{v}_-, \boldsymbol{b}) = (\boldsymbol{e}_1, w\boldsymbol{e}_2)$ for $w \in [0, 1)$, where $\boldsymbol{e}_j$ denotes the unit vector in the jth coordinate direction. Any spherically symmetric scattering map (2.2) which preserves angular momentum is thus uniquely determined by

$$\Theta(\boldsymbol{e}_1, w\boldsymbol{e}_2) = \big(\boldsymbol{e}_1 \cos\theta(w) + \boldsymbol{e}_2 \sin\theta(w), -\boldsymbol{e}_1 w \sin\theta(w) + \boldsymbol{e}_2 w \cos\theta(w)\big) \quad (2.3)$$

where $\theta(w)$ is called the *scattering angle*.

To satisfy the conditions of [21], we assume in the statements of Theorems 1.1 and 1.2 that one of the following hypotheses is true (cf. Fig. 2):

(A) $\theta \in C^1([0, 1))$ *is strictly decreasing with* $\theta(0) = \pi$ *and* $\theta(w) > 0$;
(B) $\theta \in C^1([0, 1))$ *is strictly increasing with* $\theta(0) = -\pi$ *and* $\theta(w) < 0$.

This assumption holds for a large class of scattering potentials, including muffin-tin Coulomb potentials, cf. [21]. In the case of hard-sphere scatterers we have $\theta(w) = \pi - 2\arcsin(w)$ and hence Hypothesis (A) holds. For later use we define the minimal deflection angle by

$$B_\theta := \inf_{w \in [0,1)} |\theta(w)|. \tag{2.4}$$

Note that for more general impact parameters of the form

$$b = \begin{pmatrix} 0 \\ w \end{pmatrix}, \qquad w \in \mathcal{B}_1^{d-1} \setminus \{0\}, \tag{2.5}$$

we have (by spherical symmetry)

$$\Theta\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ w \end{pmatrix}\right) = \left(S(w)\begin{pmatrix} 1 \\ 0 \end{pmatrix}, S(w)\begin{pmatrix} 0 \\ w \end{pmatrix}\right) \tag{2.6}$$

with the matrix

$$S(w) = E\big(\theta(w)\widehat{w}\big), \tag{2.7}$$

where

$$w := \|w\| > 0, \quad \widehat{w} := w^{-1}w \in S_1^{d-1}, \quad E(x) := \exp\begin{pmatrix} 0 & -{}^t x \\ x & 0_{d-1} \end{pmatrix} \in SO(d). \tag{2.8}$$

More explicitly,

$$S(w) = \begin{pmatrix} \cos\theta(w) & -{}^t\widehat{w}\sin\theta(w) \\ \widehat{w}\sin\theta(w) & 1_{d-1} - \widehat{w} \otimes \widehat{w}(1 - \cos\theta(w)) \end{pmatrix}. \tag{2.9}$$

We extend the definition of $S(w)$ to $w = 0$ by setting $S(0) := -I_d \in SO(d)$ for d even and $S(0) := \begin{pmatrix} -I_{d-1} & \\ & 1 \end{pmatrix} \in SO(d)$ for d odd. This choice ensures that $S(0)e_1 = -e_1$.

For the case of general initial data $(v_-, b) \in \mathcal{S}$, assume $R(v_-) \in SO(d)$ and $w \in \mathcal{B}_1^{d-1}$ are chosen so that

$$v_- = R(v_-)\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad b = R(v_-)\begin{pmatrix} 0 \\ w \end{pmatrix}. \tag{2.10}$$

Then

$$\Theta(v_-, b) = \left(R(v_-)S(w)\begin{pmatrix} 1 \\ 0 \end{pmatrix}, R(v_-)S(w)\begin{pmatrix} 0 \\ w \end{pmatrix}\right). \tag{2.11}$$

We use an inductive argument to work out the velocity v_n after the nth collision, as well as the impact and exit parameters b_n and s_n of the nth collision.

Lemma 2.1. *Fix v_0 and $R_0 \in SO(d)$ so that $v_0 = R_0 e_1$, and denote by $(v_n)_{n \in \mathbb{N}}$, $(b_n)_{n \in \mathbb{N}}$, $(s_n)_{n \in \mathbb{N}}$ the sequence of velocities, impact and exit parameters of a given particle trajectory. Then there is a unique sequence $(w_n)_{n \in \mathbb{N}}$ in $\mathcal{B}_1^{d-1}$ such that for all $n \in \mathbb{N}$*

$$v_n = R_n \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad b_n = R_{n-1} \begin{pmatrix} 0 \\ w_n \end{pmatrix}, \quad s_n = R_n \begin{pmatrix} 0 \\ w_n \end{pmatrix}, \tag{2.12}$$

where

$$R_n := R_0 S(w_1) \cdots S(w_n). \tag{2.13}$$

Proof. We proceed by induction. We have $v_0 \cdot b_1 = 0$ and thus $e_1 \cdot R_0^{-1} b_1 = 0$. We define $w_1 \in \mathcal{B}_1^{d-1}$ by

$$\begin{pmatrix} 0 \\ w_1 \end{pmatrix} = R_0^{-1} b_1. \tag{2.14}$$

Then the Assumption (2.10) is satisfied and (2.11) yields

$$(v_1, s_1) = \Theta(v_0, b_1) = \left(R_0 S(w_1) \begin{pmatrix} 1 \\ 0 \end{pmatrix}, R_0 S(w_1) \begin{pmatrix} 0 \\ w_1 \end{pmatrix} \right). \tag{2.15}$$

which proves the case $n = 1$. Let us therefore assume the statement is true for $n = k - 1$. By the induction hypothesis, we have $v_{k-1} = R_{k-1} e_1$. Note that $v_{k-1} \cdot b_k = 0$ implies $e_1 \cdot R_{k-1}^{-1} b_k = 0$, and define $w_k \in \mathcal{B}_1^{d-1}$ by

$$\begin{pmatrix} 0 \\ w_k \end{pmatrix} = R_{k-1}^{-1} b_k. \tag{2.16}$$

Therefore (2.10) holds with $v_- = v_{k-1}$, $b = b_k$, and we can apply (2.11):

$$\begin{aligned} (v_k, s_k) &= \Theta(v_{k-1}, b_k) \\ &= \left(R_{k-1} S(w_k) \begin{pmatrix} 1 \\ 0 \end{pmatrix}, R_{k-1} S(w_k) \begin{pmatrix} 0 \\ w_k \end{pmatrix} \right) \\ &= \left(R_k \begin{pmatrix} 1 \\ 0 \end{pmatrix}, R_k \begin{pmatrix} 0 \\ w_k \end{pmatrix} \right), \end{aligned} \tag{2.17}$$

where $R_k := R_{k-1} S(w_k) = R_0 S(w_1) \cdots S(w_k)$. This completes the proof. $\square$

3. The Boltzmann–Grad Limit

We now recall the results of [20,21] that are relevant to our investigation. Define the Markov chain

$$n \mapsto (\xi_n, \eta_n) \tag{3.1}$$

on the state space $\mathbb{R}_{>0} \times \mathcal{B}_1^{d-1}$ with transition probability

$$\mathbf{P}\left((\xi_n, \eta_n) \in \mathcal{A} \mid \xi_{n-1}, \eta_{n-1} \right) = \int_{\mathcal{A}} \Psi_0(\eta_{n-1}, x, z) \, dx \, dz. \tag{3.2}$$

We will discuss the transition kernel $\Psi_0(\boldsymbol{w}, x, z)$ in detail in Sect. 5. At this point, it suffices to note that it is independent of ξ_{n-1} and symmetric, i.e. $\Psi_0(\boldsymbol{w}, x, z) = \Psi_0(z, x, \boldsymbol{w})$. It is also independent of the choice of the scattering angle θ, the lattice $\mathcal{L}$ and the initial particle distribution Λ [20] (Note that Ψ_0 is related to the kernel Φ_0 studied in [20–22] by $\Psi_0(\boldsymbol{w}, x, z) = \Phi_0(x, \boldsymbol{w}, -z)$). Let

$$\Psi_0(x, z) := \frac{1}{v_{d-1}} \int_{\mathcal{B}_1^{d-1}} \Psi_0(\boldsymbol{w}, x, z) \, d\boldsymbol{w}, \tag{3.3}$$

$$\Psi(x, z) := \frac{1}{\overline{\xi}} \int_x^\infty \Psi_0(x', z) \, dx', \tag{3.4}$$

with the mean free path length $\overline{\xi} = 1/v_{d-1}$. Both $\Psi_0(x, z)$ and $\Psi(x, z)$ define probability densities on $\mathbb{R}_{>0} \times \mathcal{B}_1^{d-1}$ with respect to $dx \, dz$. The first fact follows from the symmetry of the transition kernel, and the second from the relation

$$\int_{\mathcal{B}_1^{d-1} \times \mathbb{R}_{>0}} \Psi(x, z) \, dx \, dz = \frac{1}{\overline{\xi}} \int_{\mathcal{B}_1^{d-1} \times \mathbb{R}_{>0}} x \Psi_0(x, z) \, dx \, dz = 1. \tag{3.5}$$

Suppose in the following that the sequence of random variables

$$\left((\xi_n, \boldsymbol{\eta}_n) \right)_{n=1}^\infty \tag{3.6}$$

is given by the Markov chain (3.1), where $(\xi_1, \boldsymbol{\eta}_1)$ has density either $\Psi(x, z)$ (for the continuous time setting) or $\Psi_0(x, z)$ (for the discrete time setting). The relation (3.4) between the two reflects the fact that the continuous time Markov process is a suspension flow over the discrete time process, where the particle moves with unit speed between consecutive collisions; see [21, Sect. 6] for more details.

We assume in the following that R is a function $S_1^{d-1} \to SO(d)$ which satisfies $\boldsymbol{v} = R(\boldsymbol{v})\boldsymbol{e}_1$ and which is smooth when restricted to $S_1^{d-1} \setminus \{-\boldsymbol{e}_1\}$. An example is

$$R(\boldsymbol{v}) = E\left(\frac{2 \arcsin(\|\boldsymbol{v} - \boldsymbol{e}_1\|/2)}{\|\boldsymbol{v}_\perp\|} \boldsymbol{v}_\perp \right) \quad \text{for} \quad \boldsymbol{v} \in S_1^{d-1} \setminus \{\boldsymbol{e}_1, -\boldsymbol{e}_1\}, \tag{3.7}$$

where $\boldsymbol{v}_\perp := (v_2, \ldots, v_d) \in \mathbb{R}^{d-1}$, and $R(\boldsymbol{e}_1) = I$, $R(-\boldsymbol{e}_1) = -I$.

For $n \in \mathbb{N}$, define the following random variables:

$$\tau_n := \sum_{j=1}^n \xi_j, \quad \tau_0 := 0, \quad \text{(time to the nth collision);} \tag{3.8}$$

$$v_t := \max\{n \in \mathbb{Z}_{\geq 0} : \tau_n \leq t\} \quad \text{(number of collisions within time t);} \tag{3.9}$$

$$\boldsymbol{V}_n := R(\boldsymbol{v}_0) S(\boldsymbol{\eta}_1) \cdots S(\boldsymbol{\eta}_n) \boldsymbol{e}_1, \quad \boldsymbol{V}_0 := \boldsymbol{v}_0, \quad \text{(velocity after the nth collision);} \tag{3.10}$$

$$\boldsymbol{Q}_n := \sum_{j=1}^n \xi_j \boldsymbol{V}_{j-1} \quad \text{(discrete time displacement);} \tag{3.11}$$

$$\boldsymbol{X}_t := \boldsymbol{Q}_{v_t} + (t - \tau_{v_t}) \boldsymbol{V}_{v_t} \quad \text{(continuous time displacement).} \tag{3.12}$$

Theorem 3.1 ([21]).

(i) *Under the hypotheses of Theorem 1.1, for any $t > 0$,*

$$x_t - x_0 \Rightarrow X_t \tag{3.13}$$

as $r \to 0$, where the random variable (ξ_1, η_1) has density $\Psi(x, z)$.

(ii) *Under the hypotheses of Theorem 1.2, for any $n \in \mathbb{N}$,*

$$q_n - q_0 \Rightarrow Q_n \tag{3.14}$$

as $r \to 0$, where the random variable (ξ_1, η_1) has density $\Psi_0(x, z)$.

The main part of this paper is devoted to the proof of the following superdiffusive central limit theorem for the processes X_t and Q_n, which in turn implies Theorems 1.1 and 1.2. We will only assume that the random variable (ξ_1, η_1) is such that the marginal distribution of η_1 is absolutely continuous on $\mathcal{B}_1^{d-1}$ with respect to Lebesgue measure; there is no further assumption on the distribution of ξ_1. This hypothesis is satisfied for (ξ_1, η_1) with density $\Psi_0(x, z)$, since

$$\overline{\Psi}_0(z) := \int_0^\infty \Psi_0(x, z)\, dx = \frac{1}{v_{d-1}} \int_{\mathbb{R}_{>0} \times \mathcal{B}_1^{d-1}} \Psi_0(z, x, w)\, dx\, dw = \frac{1}{v_{d-1}}. \tag{3.15}$$

That is, the marginal distribution of η_1 is uniform on $\mathcal{B}_1^{d-1}$. We will later see that (ξ_1, η_1) with density $\Psi(x, z)$ also complies with the above hypothesis (cf. Proposition 10.1). The processes X_t and Q_n are independent of x_0 and q_0, respectively, and we will in the following fix $v_0 \in S_1^{d-1}$. Also, the required assumptions on the scattering angle θ are significantly weaker than in the previous theorems.

Theorem 3.2. *Let $d \geq 2$, $v_0 \in S_1^{d-1}$ and assume that the marginal distribution of η_1 is absolutely continuous. Assume $\theta : [0, 1) \to [-\pi, \pi]$ is measurable, so that*

$$\text{meas}\{w \in [0, 1) : \theta(w) \notin \pi\mathbb{Q}\} > 0. \tag{3.16}$$

Then (i)

$$\frac{X_t}{\Sigma_d \sqrt{t \log t}} \Rightarrow \mathcal{N}(0, I_d), \tag{3.17}$$

and (ii)

$$\frac{Q_n}{\sigma_d \sqrt{n \log n}} \Rightarrow \mathcal{N}(0, I_d). \tag{3.18}$$

In view of Theorem 3.1, Theorem 3.2 implies Theorems 1.1 and 1.2.

Statement (ii) in Theorem 3.1 generalizes to the convergence of the random curve (3.1) obtained by linearly interpolating q_n [21, Theorem 1.1]. That is, under the conditions of Theorem 1.4, for $r \to 0$ and arbitrary fixed n,

$$Y_{n,r} \Rightarrow Y_n \tag{3.19}$$

where the rescaled discrete-time limiting process is defined by

$$Y_n(t) := \frac{Q_n(t)}{\sigma_d \sqrt{n \log n}}, \tag{3.20}$$

and

$$\boldsymbol{Q}_n(t) := \boldsymbol{Q}_{\lfloor nt \rfloor} + \{nt\}\, \xi_{\lfloor nt \rfloor + 1} \boldsymbol{V}_{\lfloor nt \rfloor} \tag{3.21}$$

denotes the linear interpolation of the discrete time displacements $\boldsymbol{Q}_0, \boldsymbol{Q}_1, \boldsymbol{Q}_2, \ldots$ We will prove in Sect. 12 that the $\boldsymbol{Y}_n$ converges to $\boldsymbol{W}$ in finite-dimensional distribution. The last missing ingredient in the proof of Theorem 1.4 is then the tightness of the probability measures associated with the sequence of random curves $(\boldsymbol{Y}_n)_{n=1}^{\infty}$ in $C_0([0, 1])$, which is established in Sect. 13. Theorem 1.3 follows from Theorem 1.4 via estimates presented in Sect. 14.

It is interesting to compare the above results with the case of a random, rather than periodic, scatterer configuration, where the scatterers are placed at the points of a fixed realisation of a Poisson process in $\mathbb{R}^d$. In the case of fixed scattering radius there is, to the best of our knowledge, no proof of a central limit theorem even in dimension $d = 2$. In the Boltzmann–Grad limit, however, the work of Gallavotti [16], Spohn [26] and Boldrighini et al. [7] shows that we have an analogue of Theorem 3.1, where the limit random flight process X_t is governed by the linear Boltzmann equation. In this setting, (3.6) is a sequence of independent random variables, where ξ_n has density $\Psi_0(x) = v_{d-1} \exp(-v_{d-1}x)$ and $\boldsymbol{\eta}_n$ is uniformly distributed in $\mathcal{B}_1^{d-1}$. Routine techniques [25] show that in this case the central limit theorem holds for X_t with a standard $\sqrt{t}$ normalisation, and for $\boldsymbol{Q}_n$ with a $\sqrt{n}$ normalisation.

4. Outline of the Proof of Theorem 3.2

We will now outline the central arguments in the proof of Theorem 3.2 (ii) for discrete time by reducing the statement to four main lemmas, whose proof is given in Sect. 9. The continuous-time case (i) follows from (ii) via technical estimates supplied in Sect. 11. We will assume from now on that $(\xi_1, \boldsymbol{\eta}_1)$ has density $\Psi_0(x, z)$, and discuss the generalisation to more general distributions in Sect. 10. We note that for $\boldsymbol{\eta}_0$ uniformly distributed in $\mathcal{B}_1^{d-1}$,

$$\Psi_0(x, z) = \mathbf{E}\, \Psi_0(\boldsymbol{\eta}_0, x, z), \tag{4.1}$$

and it is therefore equivalent to consider instead of (3.6) the Markov chain

$$\big((\xi_n, \boldsymbol{\eta}_n)\big)_{n=0}^{\infty} \tag{4.2}$$

with the same transition probability (3.2), $\boldsymbol{\eta}_0$ uniformly distributed in $\mathcal{B}_1^{d-1}$ and $\xi_0 = 0$. The sequence

$$\underline{\boldsymbol{\eta}} = (\boldsymbol{\eta}_n)_{n=0}^{\infty}, \tag{4.3}$$

with $\boldsymbol{\eta}_0$ as defined above, is itself generated by a Markov chain on the state space $\mathcal{B}_1^{d-1}$ with transition probability

$$\mathbf{P}\big(\boldsymbol{\eta}_n \in \mathcal{A} \mid \boldsymbol{\eta}_{n-1}\big) = \int_{\mathcal{A}} K_0(\boldsymbol{\eta}_{n-1}, z)\, dz \tag{4.4}$$

where

$$K_0(\boldsymbol{w}, z) := \int_0^{\infty} \Psi_0(\boldsymbol{w}, x, z)\, dx. \tag{4.5}$$

The objective is to prove a central limit theorem of sums of the random variables $\xi_n V_{n-1}$. The first observation is that these are of course not independent. If we, however, condition on the sequence $\underline{\eta}$, then the V_n are deterministic, and $(\xi_n)_{n=1}^{\infty}$ is a sequence of independent (but not identically distributed) random variables,

$$\mathbf{P}\left(\xi_n \in (x, x + dx) \mid \underline{\eta}\right) = \frac{\Psi_0(\eta_{n-1}, x, \eta_n)\, dx}{K_0(\eta_{n-1}, \eta_n)}. \tag{4.6}$$

The plan is now to apply the Lindeberg central limit theorem to the sum of independent random variables, $Q_n = \sum_{j=1}^{n} \xi_j V_{j-1}$, conditioned on $\underline{\eta}$.

To this end we first truncate Q_n by defining the random variable

$$Q'_n := \sum_{j=1}^{n} \xi'_j V_{j-1} \tag{4.7}$$

with

$$\xi'_j := \xi_j \mathbb{1}_{\{\xi_j^2 \leq j(\log j)^\gamma\}} \tag{4.8}$$

for some fixed $\gamma \in (1, 2)$. The following lemma tells us that it is sufficient to prove Theorem 3.2 (ii) for Q'_n instead of Q_n.

Lemma 4.1. *We have*

$$\sup_{n \in \mathbb{N}} \| Q_n - Q'_n \| < \infty \tag{4.9}$$

almost surely.

To prove the central limit theorem for Q'_n, we center ξ'_j by setting

$$\tilde{\xi}_j = \xi'_j - m_j, \tag{4.10}$$

with the conditional expectation

$$m_j := \mathbf{E}\left(\xi'_j \mid \underline{\eta}\right) = \frac{K_{1,r_j}(\eta_{j-1}, \eta_j)}{K_0(\eta_{j-1}, \eta_j)} \tag{4.11}$$

where $r_j := \sqrt{j(\log j)^\gamma}$ and

$$K_{1,r}(w, z) := \int_0^r x \Psi_0(w, x, z)dx. \tag{4.12}$$

Let

$$\tilde{Q}_n := \sum_{j=1}^{n} \tilde{\xi}_j V_{j-1}. \tag{4.13}$$

The following lemma shows that Q'_n and $\tilde{Q}_n$ are close relative to $\sqrt{n \log n}$.

Lemma 4.2. *The sequence of random variables*

$$\frac{Q'_n - \widetilde{Q}_n}{\sqrt{n \log \log n}} \tag{4.14}$$

is tight if $d = 2$, and

$$\frac{Q'_n - \widetilde{Q}_n}{\sqrt{n}} \tag{4.15}$$

is tight if $d \geq 3$.

It is therefore sufficient to prove Theorem 3.2 (ii) for $\widetilde{Q}_n$ in place of Q_n. This will be achieved by applying the Lindeberg central limit theorem to the conditional sum as alluded to above. We begin by estimating the conditional variance. Set

$$a_j^2 := \mathbf{Var}\left(\tilde{\xi}_j \mid \underline{\eta}\right) = \frac{K_{2,r_j}(\eta_{j-1}, \eta_j)}{K_0(\eta_{j-1}, \eta_j)} - m_j^2, \tag{4.16}$$

with

$$K_{2,r}(\boldsymbol{w}, \boldsymbol{z}) := \int_0^r x^2 \Psi_0(\boldsymbol{w}, x, \boldsymbol{z}) dx. \tag{4.17}$$

Lemma 4.3. *There is a constant $\sigma_d > 0$ such that, for $n \to \infty$,*

$$\frac{\mathbf{E}\left(\widetilde{Q}_n \otimes \widetilde{Q}_n \mid \underline{\eta}\right)}{n \log n} = \frac{\sum_{j=1}^n a_j^2 V_{j-1} \otimes V_{j-1}}{n \log n} \xrightarrow{\mathbf{P}} \sigma_d^2 I_d. \tag{4.18}$$

By taking the trace in (4.18), we have in particular

$$\frac{A_n^2}{n \log n} \xrightarrow{\mathbf{P}} d\,\sigma_d^2 \tag{4.19}$$

for

$$A_n^2 := \sum_{j=1}^n a_j^2 = \mathbf{E}\left(\|\widetilde{Q}_n\|^2 \mid \underline{\eta}\right). \tag{4.20}$$

Recall that convergence in probability $x_n \xrightarrow{\mathbf{P}} x$ is defined as $\lim_{n \to \infty} \mathbf{P}(|x_n - x| > \epsilon) = 0$ for any $\epsilon > 0$.

The next lemma verifies the Lindeberg conditions for random $\underline{\eta}$.

Lemma 4.4. *For any fixed $\varepsilon > 0$,*

$$A_n^{-2} \sum_{j=1}^n \mathbf{E}\left(\tilde{\xi}_j^2 \mathbb{1}_{\{\tilde{\xi}_j^2 > \varepsilon^2 A_n^2\}} \mid \underline{\eta}\right) \xrightarrow{\mathbf{P}} 0 \tag{4.21}$$

as $n \to \infty$.

Given these lemmas, let us now conclude the proof of the fact that

$$Y_n := \frac{\widetilde{Q}_n}{\sigma_d \sqrt{n \log n}} \Rightarrow \mathcal{N}(0, I_d). \tag{4.22}$$

By Chebyshev's inequality we have, for any $K > 0$,

$$\mathbf{P}\left(\|Y_n\| > K \mid \underline{\eta}\right) \leq \frac{1}{K^2} \mathbf{E}\left(\|Y_n\|^2 \mid \underline{\eta}\right), \tag{4.23}$$

and thus, for any $\kappa > 0$,

$$\begin{aligned}
\mathbf{P}\left(\|Y_n\| > K\right) &\leq \frac{\kappa^2}{K^2} + \mathbf{P}\left(\mathbf{E}\left(\|Y_n\|^2 \mid \underline{\eta}\right) > \kappa^2\right) \\
&= \frac{\kappa^2}{K^2} + \mathbf{P}\left(A_n^2 > \kappa^2 \sigma_d^2 \, n \log n\right).
\end{aligned} \tag{4.24}$$

By (4.19), the second term on the right hand side of (4.24) converges to 0 as $n \to \infty$, if we choose $\kappa = d$, say. So (4.24) implies that the sequence of random variables Y_n is tight. By the Helly–Prokhorov theorem, there is an infinite subset $S_1 \subset \mathbb{N}$ so that Y_n converges in distribution along $n \in S_1$ to some limit Y. Assume for a contradiction that Y is *not* distributed according to $\mathcal{N}(0, I_d)$. The Borel–Cantelli lemma implies that there is an infinite subset $S_2 \subset S_1$, so that in the statements of Lemmas 4.3 and 4.4 we have almost-sure convergence along $n \in S_2$:

$$\frac{\mathbf{E}\left(\widetilde{Q}_n \otimes \widetilde{Q}_n \mid \underline{\eta}\right)}{n \log n} \xrightarrow{\text{a.s.}} \sigma_d^2 \, I_d, \tag{4.25}$$

$$\frac{A_n^2}{n \log n} \xrightarrow{\text{a.s.}} d \, \sigma_d^2, \tag{4.26}$$

and

$$A_n^{-2} \sum_{j=1}^{n} \mathbf{E}\left(\tilde{\xi}_j^2 \mathbb{1}_{\{\tilde{\xi}_j^2 > \varepsilon^2 A_n^2\}} \mid \underline{\eta}\right) \xrightarrow{\text{a.s.}} 0. \tag{4.27}$$

The hypotheses of the Lindeberg central limit theorem are met, and we infer that $Y_n \Rightarrow \mathcal{N}(0, I_d)$ for $n \in S_2$. (We use the Lindeberg theorem for *triangular arrays* of independent random variables, since we have verified the Lindeberg conditions only along a subsequence.) This, however, contradicts our assumption that Y is not normal, and hence $\mathcal{N}(0, I_d)$ is indeed the unique limit point of any converging subsequence. This in turn implies that every sequence converges, and therefore completes the proof of (4.22). In view of Lemmas 4.1 and 4.2, this implies Theorem 3.2 (ii) (still under the assumption that (ξ_1, η_1) has density $\Psi_0(x, z)$).

Let us briefly describe the further contents of this paper. In Sect. 5 we recall the basic properties of the transition kernel $\Psi_0(w, x, z)$ from [22]. Section 6 establishes key estimates for the moments $K_{p,r}(w, z)$, m_j and a_j introduced above. In Sects. 7 and 8 we prove spectral gap estimates and exponential mixing for the discrete time Markov process defined in (4.4). The estimates from Sects. 6–8 are the main input in the proof of Lemmas 4.1–4.4, which is given in Sect. 9. In Sect. 10 we show that the discrete-time statement in Theorem 3.2 (ii) holds for more general initial distributions than $\Psi_0(x, z)$. It holds in particular for $\Psi(x, z)$, which appears in the continuous-time variant. Section 11 explains how to pass from discrete to continuous time, thus completing the proof of Theorem 3.2 (i).

5. The Transition Kernel

In dimension $d = 2$ we have the following explicit formula for the transition kernel. For $w, z \in (-1, 1)$,

$$\Psi_0(w, x, z) = \frac{6}{\pi^2} \begin{cases} \mathbb{1}_{\{0 \leq x < \frac{1}{1+w}\}} + \mathbb{1}_{\{\frac{1}{1+w} \leq x < \frac{1}{1+z}\}} \dfrac{x^{-1} - 1 - z}{w - z} & \text{if } 0 < w, \ -w < z < w, \\[2ex] \mathbb{1}_{\{0 \leq x < \frac{1}{1+z}\}} + \mathbb{1}_{\{\frac{1}{1+z} \leq x < \frac{1}{1+w}\}} \dfrac{x^{-1} - 1 - w}{z - w} & \text{if } 0 < z, \ -z < w < z, \\[2ex] \mathbb{1}_{\{0 \leq x < \frac{1}{1-w}\}} + \mathbb{1}_{\{\frac{1}{1-w} \leq x < \frac{1}{1-z}\}} \dfrac{x^{-1} - 1 + z}{z - w} & \text{if } w < 0, \ w < z < -w, \\[2ex] \mathbb{1}_{\{0 \leq x < \frac{1}{1-z}\}} + \mathbb{1}_{\{\frac{1}{1-z} \leq x < \frac{1}{1-w}\}} \dfrac{x^{-1} - 1 + w}{w - z} & \text{if } z < 0, \ z < w < -z. \end{cases}$$
(5.1)

This formula has been derived, independently and with different methods, by Marklof and Strömbergsson [19], Caglioti and Golse [11,12] and by Bykovskii and Ustinov [10].

In dimension $d \geq 3$ we have no such explicit formulas for the transition kernel. We recall from [20–22] the following properties. If $d \geq 3$, the function

$$\Psi_0 : \mathcal{B}_1^{d-1} \times \mathbb{R}_{>0} \times \mathcal{B}_1^{d-1} \to [0, 1]$$
(5.2)

is continuous. $\Psi_0(\boldsymbol{w}, x, \boldsymbol{z})$ depends only on x, $w := \|\boldsymbol{w}\|$, $z := \|\boldsymbol{z}\|$ and the angle $\varphi := \varphi(\boldsymbol{w}, \boldsymbol{z}) \in [0, \pi]$ between the vectors $\boldsymbol{w}, \boldsymbol{z} \in \mathcal{B}_1^{d-1}$. Note that in dimension $d = 2$ the angle φ can only take the values 0 and π. For statements that are specific to dimension $d = 2$, we will often use $w \in (-1, 1)$ instead of $\boldsymbol{w}$, and $|w|$ instead of $w = \|\boldsymbol{w}\|$. We recall once more that $\Psi_0(\boldsymbol{w}, x, \boldsymbol{z}) = \Phi_0(x, \boldsymbol{w}, -\boldsymbol{z})$ in the notation of [20–22], and so in particular the angle φ between $\boldsymbol{w}, \boldsymbol{z}$ becomes $\pi - \varphi$.

Our proofs will exploit the following estimates on the transition kernel [22,28]. All bounds are uniform in $x > 0$ and $\boldsymbol{w}, \boldsymbol{z} \in \mathcal{B}_1^{d-1}$. We have by [22, Theorem 1.1],

$$\frac{1 - 2^{d-1} v_{d-1} x}{\zeta(d)} \leq \Psi_0(\boldsymbol{w}, x, \boldsymbol{z}) \leq \frac{1}{\zeta(d)}.$$
(5.3)

Furthermore, by [22, Theorem 1.7], there exists a continuous and uniformly bounded function $F_{0,d} : \mathbb{R}_{>0} \times \mathbb{R}_{>0} \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ such that

$$\Psi_0(\boldsymbol{w}, x, \boldsymbol{z}) = x^{-2+\frac{2}{d}} F_{0,d}\left(x^{\frac{2}{d}}(1-z), x^{\frac{2}{d}}(1-w), x^{\frac{1}{d}}(\pi - \varphi)\right) + O(E), \quad (5.4)$$

where the error term is

$$E = \begin{cases} x^{-2} & \text{if } d = 2, \\ x^{-2} \log(2 + \min(x, (\pi - \varphi)^{-1})) & \text{if } d = 3, \\ \min\left(x^{-2}, x^{-3+\frac{2}{d-1}}(\pi - \varphi)^{2-d+\frac{2}{d-1}}\right) & \text{if } d \geq 4. \end{cases}$$
(5.5)

It is noted in [22] that $F_{0,d}(t_1, t_2, \alpha)$ is uniformly bounded from below for t_1, t_2, α near zero. That is, there is a small constant $c > 0$ which only depends on d such that

$$\max(t_1, t_2, \alpha) < c \implies F_{0,d}(t_1, t_2, \alpha) > c.$$
(5.6)

Furthermore, the support of $F_{0,d}$ is contained in $(0, c'] \times (0, c'] \times \mathbb{R}_{\geq 0}$ for some $c' > 0$, and for any fixed $t_1, t_2 > 0$, the function $F_{0,d}(t_1, t_2, \cdot)$ has compact support.

In dimension $d \geq 3$, the following upper bound will prove useful [28, Theorem 1.8]:

$$\Psi_0(\boldsymbol{w}, x, z) \ll \begin{cases} x^{-2} \min\left\{1, (x\varphi^{d-2})^{-1+\frac{2}{d-1}}\right\} & \text{if } \varphi \in [0, \frac{\pi}{2}] \\ x^{-2+\frac{2}{d}} \min\left\{1, (x(\pi - \varphi)^d)^{-1+\frac{2}{d(d-1)}}\right\} & \text{if } \varphi \in [\frac{\pi}{2}, \pi]. \end{cases} \tag{5.7}$$

The notation $f \ll g$ is here defined as $f = O(g)$, i.e., there exists a constant $C > 0$ such that $|f| \leq C|g|$. The notation $f \asymp g$ used below means that $g \ll f \ll g$, i.e., there exist a constant $C \geq 1$ such that $C^{-1}|g| \leq |f| \leq C|g|$.

The support of $\Psi_0(\boldsymbol{w}, x, z)$ is described by a continuous function $x_0 : \mathcal{B}_1^{d-1} \times \mathcal{B}_1^{d-1} \to \mathbb{R}_{>0}$. We have $\Psi_0(\boldsymbol{w}, x, z) > 0$ holds if and only if $x < x_0(\boldsymbol{w}, z)$. Set

$$t := t(\boldsymbol{w}, z) := \max(1 - w, 1 - z) \in (0, 1]. \tag{5.8}$$

If $d \geq 3$, then [28, Proposition 1.9] tells us that

$$x_0(\boldsymbol{w}, z) \asymp \begin{cases} \max(t^{-\frac{d-2}{2}}, t^{-\frac{d-1}{2}}\varphi) & \text{if } \varphi \in [0, \frac{\pi}{2}] \\ \min(t^{-\frac{d}{2}}, t^{-\frac{d-1}{2}}(\pi - \varphi)^{-1}) & \text{if } \varphi \in [\frac{\pi}{2}, \pi]. \end{cases} \tag{5.9}$$

(If $\varphi = \pi$ then the right hand side of (5.9) should be interpreted as $t^{-\frac{d}{2}}$.)

For the distribution of free path length between consecutive collisions,

$$\Psi_0(x) = \frac{1}{v_{d-1}} \int_{\mathcal{B}_1^{d-1}} \int_{\mathcal{B}_1^{d-1}} \Psi_0(\boldsymbol{w}, x, z) \, d\boldsymbol{w} \, dz, \tag{5.10}$$

we have the following tail estimate [22, Theorem 1.14] (Ψ_0 is denoted $\overline{\Phi}_0$ in [22]): For $x \to \infty$,

$$\Psi_0(x) = \Theta_d \, x^{-3} + O\left(x^{-3-\frac{2}{d}}\right) \times \begin{cases} 1 & \text{if } d = 2 \\ \log x & \text{if } d = 3 \\ 1 & \text{if } d \geq 4 \end{cases} \tag{5.11}$$

with

$$\Theta_d := \frac{2^{2-d}}{d(d+1)\zeta(d)}. \tag{5.12}$$

This asymptotic estimate sharpens earlier upper and lower bounds by Bourgain et al. [8,17]. Note that the variances in Theorems 1.1 and 1.2 are related to the above tail via

$$\Sigma_d^2 = \frac{\Theta_d}{2d\bar{\xi}}, \quad \sigma_d^2 = \frac{\Theta_d}{2d}. \tag{5.13}$$

6. Moment Estimates

We now provide key estimates of the random variables introduced in the previous section. For $p = 0, 1, 2$ and $r > 0$, set

$$K_{p,r}(\boldsymbol{w}, \boldsymbol{z}) := \int_0^r x^p \Psi_0(\boldsymbol{w}, x, \boldsymbol{z}) dx \tag{6.1}$$

and

$$K_p(\boldsymbol{w}, \boldsymbol{z}) := \int_0^\infty x^p \Psi_0(\boldsymbol{w}, x, \boldsymbol{z}) dx. \tag{6.2}$$

Note that $v_{d-1}^{-1} K_0(\boldsymbol{w}, \boldsymbol{z}) \, d\boldsymbol{w} \, d\boldsymbol{z}$ defines a probability measure on $\mathcal{B}_1^{d-1} \times \mathcal{B}_1^{d-1}$. We furthermore define the random variables (recall Sect. 4)

$$m_j := \mathbf{E}\left(\xi_j' \mid \underline{\boldsymbol{\eta}}\right) = \frac{K_{1,r_j}(\boldsymbol{\eta}_{j-1}, \boldsymbol{\eta}_j)}{K_0(\boldsymbol{\eta}_{j-1}, \boldsymbol{\eta}_j)}, \quad \mu_j := \mathbf{E}\left(\xi_j \mid \underline{\boldsymbol{\eta}}\right) = \frac{K_1(\boldsymbol{\eta}_{j-1}, \boldsymbol{\eta}_j)}{K_0(\boldsymbol{\eta}_{j-1}, \boldsymbol{\eta}_j)}, \tag{6.3}$$

$$b_j^2 := \mathbf{E}\left(\xi_j'^2 \mid \underline{\boldsymbol{\eta}}\right) = \frac{K_{2,r_j}(\boldsymbol{\eta}_{j-1}, \boldsymbol{\eta}_j)}{K_0(\boldsymbol{\eta}_{j-1}, \boldsymbol{\eta}_j)}, \quad \beta_j^2 := \mathbf{E}\left(\xi_j^2 \mid \underline{\boldsymbol{\eta}}\right) = \frac{K_2(\boldsymbol{\eta}_{j-1}, \boldsymbol{\eta}_j)}{K_0(\boldsymbol{\eta}_{j-1}, \boldsymbol{\eta}_j)}, \tag{6.4}$$

$$a_j^2 := \mathbf{Var}\left(\xi_j' \mid \underline{\boldsymbol{\eta}}\right) = b_j^2 - m_j^2, \quad \alpha_j^2 := \mathbf{Var}\left(\xi_j \mid \underline{\boldsymbol{\eta}}\right) = \beta_j^2 - \mu_j^2, \tag{6.5}$$

and

$$A_n^2 := \sum_{j=1}^n a_j^2, \tag{6.6}$$

with $r_j = \sqrt{j(\log j)^\gamma}$ for some fixed $\gamma \in (1, 2)$.

Lemma 6.1. *Let $d = 2$. Then, for $w, z \in (-1, 1)$,*

$$K_0(w, z) = \frac{6}{\pi^2} \begin{cases} \dfrac{1}{w-z} \ln \dfrac{1+w}{1+z} & \text{if } w + z \geq 0, \\[2mm] \dfrac{1}{z-w} \ln \dfrac{1-w}{1-z} & \text{if } w + z \leq 0, \end{cases} \tag{6.7}$$

$$K_1(w, z) = \frac{3}{\pi^2} \begin{cases} \dfrac{1}{(1+w)(1+z)} & \text{if } w + z \geq 0, \\[2mm] \dfrac{1}{(1-w)(1-z)} & \text{if } w + z \leq 0, \end{cases} \tag{6.8}$$

$$K_2(w, z) = \frac{1}{\pi^2} \begin{cases} \dfrac{2+w+z}{(1+w)^2(1+z)^2} & \text{if } w + z \geq 0, \\[2mm] \dfrac{2-w-z}{(1-w)^2(1-z)^2} & \text{if } w + z \leq 0. \end{cases} \tag{6.9}$$

Proof. These follow from the explicit formula (5.1) by direct computation. □

Lemma 6.2. *Let $d \geq 3$. Then*

$$\inf_{\boldsymbol{w}, \boldsymbol{z} \in \mathcal{B}_1^{d-1}} K_0(\boldsymbol{w}, \boldsymbol{z}) \geq \frac{1}{2^d v_{d-1} \zeta(d)} > 0, \tag{6.10}$$

$$\sup_{\boldsymbol{w}, \boldsymbol{z} \in \mathcal{B}_1^{d-1}} K_0(\boldsymbol{w}, \boldsymbol{z}) < \infty. \tag{6.11}$$

Proof. We have

$$K_0(\boldsymbol{w}, \boldsymbol{z}) \geq \int_0^y \Psi_0(\boldsymbol{w}, x, \boldsymbol{z}) dx \tag{6.12}$$

for any $y \geq 0$. Theorem 1.1 in [22] states that for $x > 0$ and $\boldsymbol{w}, \boldsymbol{z} \in \mathcal{B}_1^{d-1}$,

$$\Psi_0(\boldsymbol{w}, x, \boldsymbol{z}) \geq \frac{1 - 2^{d-1} v_{d-1} x}{\zeta(d)}, \tag{6.13}$$

and the lower bound follows with the choice $y = (2^{d-1} v_{d-1})^{-1}$. The upper bound follows from (5.7) which tells us that

$$\Psi_0(\boldsymbol{w}, x, \boldsymbol{z}) = O(x^{-2+\frac{2}{d}}), \tag{6.14}$$

where the implied constant is independent of $x, \boldsymbol{w}, \boldsymbol{z}$. □

Lemma 6.3. *Let $d \geq 3$. For $\boldsymbol{w}, \boldsymbol{z} \in \mathcal{B}_1^{d-1}$,*

$$K_1(\boldsymbol{w}, \boldsymbol{z}) \gg \min(t^{-1}, (\pi - \varphi)^{-2}), \tag{6.15}$$

and

$$K_1(\boldsymbol{w}, \boldsymbol{z}) \ll \begin{cases} 1 + \log \max(t^{-\frac{d-2}{2}}, t^{-\frac{d-1}{2}} \varphi) & \text{if } \varphi \in [0, \frac{\pi}{2}], \\ \min(t^{-1}, t^{-1+\frac{1}{d}} (\pi - \varphi)^{-\frac{2}{d}}) & \text{if } \varphi \in [\frac{\pi}{2}, \pi]. \end{cases} \tag{6.16}$$

Proof. As to the lower bound (6.15), we note that by (5.4)

$$K_1(\boldsymbol{w}, \boldsymbol{z})$$
$$\geq \int_{\substack{x>0 \\ x^{\frac{2}{d}} t(\boldsymbol{w},\boldsymbol{z})<c \\ x^{\frac{1}{d}}(\pi-\varphi)<c}} x \left\{ x^{-2+\frac{2}{d}} F_{0,d}\left(x^{\frac{2}{d}}(1-z), x^{\frac{2}{d}}(1-w), x^{\frac{1}{d}}(\pi-\varphi)\right) - |O(E)| \right\} dx,$$
$$\tag{6.17}$$

which, in view of (5.5) and (5.6), implies

$$K_1(\boldsymbol{w}, \boldsymbol{z}) > c \int_{\substack{x>0 \\ x^{\frac{2}{d}} t(\boldsymbol{w},\boldsymbol{z})<c \\ x^{\frac{1}{d}}(\pi-\varphi)<c}} \left\{ x^{-1+\frac{2}{d}} - |O((1+x)^{-1} \log(2+x))| \right\} dx \tag{6.18}$$
$$\gg \min\left(t(\boldsymbol{w}, \boldsymbol{z})^{-1}, (\pi - \varphi)^{-2} \right).$$

The upper bound (6.16) follows from (5.7) and (5.9): for $\varphi \in [0, \frac{\pi}{2}]$,

$$K_1(\boldsymbol{w}, \boldsymbol{z}) \ll 1 + \int_1^{x_0(\boldsymbol{w},\boldsymbol{z})} x^{-1} dx \ll 1 + \log \max(t^{-\frac{d-2}{2}}, t^{-\frac{d-1}{2}} \varphi), \tag{6.19}$$

and for $\varphi \in [\frac{\pi}{2}, \pi]$, we have

$$K_1(\boldsymbol{w}, \boldsymbol{z}) \ll \int_0^{x_0(\boldsymbol{w},\boldsymbol{z})} x^{-1+\frac{2}{d}} dx \ll \min(t^{-1}, t^{-1+\frac{1}{d}} (\pi - \varphi)^{-\frac{2}{d}}). \tag{6.20}$$

Lemma 6.4. *Let $d \geq 3$. Then, for $\mathbf{w}, \mathbf{z} \in \mathcal{B}_1^{d-1}$,*

$$K_2(\mathbf{w}, \mathbf{z}) \gg \min(t^{-(1+\frac{d}{2})}, (\pi - \varphi)^{-(d+2)}), \tag{6.21}$$

and

$$K_2(\mathbf{w}, \mathbf{z}) \ll \begin{cases} \max(t^{-\frac{d-2}{2}}, t^{-\frac{d-1}{2}}\varphi) & \text{if } \varphi \in [0, \frac{\pi}{2}], \\ \min(t^{-(1+\frac{d}{2})}, t^{-\frac{d+1}{2}+\frac{1}{d}}(\pi - \varphi)^{-(1+\frac{2}{d})}) & \text{if } \varphi \in [\frac{\pi}{2}, \pi]. \end{cases} \tag{6.22}$$

Proof. The lower bound (6.21) follows from

$$K_2(\mathbf{w}, \mathbf{z})$$
$$\geq \int_{\substack{x > 0 \\ x^{\frac{2}{d}}t(\mathbf{w},\mathbf{z}) < c \\ x^{\frac{1}{d}}(\pi-\varphi) < c}} x^2 \left\{ x^{-2+\frac{2}{d}} F_{0,d}\left(x^{\frac{2}{d}}(1-z), x^{\frac{2}{d}}(1-w), x^{\frac{1}{d}}(\pi - \varphi)\right) - |O(E)| \right\} dx, \tag{6.23}$$

and (5.5) and (5.6). Hence

$$K_2(\mathbf{w}, \mathbf{z}) > c \int_{\substack{x > 0 \\ x^{\frac{2}{d}}t(\mathbf{w},\mathbf{z}) < c \\ x^{\frac{1}{d}}(\pi-\varphi) < c}} \left\{ x^{\frac{2}{d}} - |O(\log(2 + x))| \right\} dx \tag{6.24}$$

$$\gg \min\left(t(\mathbf{w}, \mathbf{z})^{-(1+\frac{d}{2})}, (\pi - \varphi)^{-(d+2)} \right).$$

The upper bound (6.22) follows from (5.7) and (5.9): for $\varphi \in [0, \frac{\pi}{2}]$,

$$K_2(\mathbf{w}, \mathbf{z}) \ll \int_0^{x_0(\mathbf{w},\mathbf{z})} dx \ll \max(t^{-\frac{d-2}{2}}, t^{-\frac{d-1}{2}}\varphi), \tag{6.25}$$

and for $\varphi \in [\frac{\pi}{2}, \pi]$, we have

$$K_2(\mathbf{w}, \mathbf{z}) \ll \int_0^{x_0(\mathbf{w},\mathbf{z})} x^{\frac{2}{d}} dx \ll \min(t^{-(1+\frac{d}{2})}, t^{-\frac{d+1}{2}+\frac{1}{d}}(\pi - \varphi)^{-(1+\frac{2}{d})}). \tag{6.26}$$

$\square$

Proposition 6.5. *Let $d = 2$. For $u \to \infty$,*

$$\mathbf{P}(\mu_j > u) \sim \frac{3}{4\pi^2} \frac{1}{u^2 \log u}. \tag{6.27}$$

Proof. By the invariance of the integrand under $(w, z) \mapsto -(w, z)$ and $(w, z) \mapsto (z, w)$ we have

$$\mathbf{P}(\mu_j > u) = \frac{1}{2} \int_{(-1,1)^2} \mathbb{1}_{\{K_1(w,z) > u K_0(w,z)\}} K_0(w, z) \, dw \, dz$$
$$= 2 \int_{|z| < w < 1} \mathbb{1}_{\{K_1(w,z) > u K_0(w,z)\}} K_0(w, z) \, dw \, dz. \tag{6.28}$$

(The factor $\frac{1}{2} = v_1^{-1}$ corresponds to the normalization in the remark following (6.2)). In this range of integration, we have explicitly

$$K_0(w, z) = \frac{6}{\pi^2} \frac{1}{w - z} \ln \frac{1 + w}{1 + z} \qquad (6.29)$$

and

$$\frac{K_0(w, z)}{K_1(w, z)} = 2 \frac{(1 + w)(1 + z)}{w - z} \ln \frac{1 + w}{1 + z}. \qquad (6.30)$$

Using the variable substitution

$$x = w - z, \quad y = \frac{1 + z}{w - z}, \qquad (6.31)$$

we have, with the shorthand $f(y) = 2y(1 + y) \ln(1 + y^{-1})$,

$$\begin{aligned}
\mathbf{P}(\mu_j > u) = &\frac{12}{\pi^2} \int_0^\infty \int_{\frac{2}{1+2y}}^{\frac{2}{1+y}} \mathbb{1}_{\{xf(y)<u^{-1}\}} \ln(1 + y^{-1}) \, dx \, dy \\
= &\frac{12}{\pi^2} \int_0^\infty \mathbb{1}_{\{\frac{2f(y)}{1+y} <u^{-1}\}} \left(\frac{2}{1 + y} - \frac{2}{1 + 2y} \right) \ln(1 + y^{-1}) \, dy \qquad (6.32) \\
&+ \frac{12}{\pi^2} \int_0^\infty \mathbb{1}_{\{\frac{2f(y)}{1+2y} <u^{-1}< \frac{2f(y)}{1+y}\}} \left(\frac{1}{uf(y)} - \frac{2}{1 + 2y} \right) \ln(1 + y^{-1}) \, dy.
\end{aligned}$$

The first term equals

$$\frac{12}{\pi^2} \int_0^\infty \mathbb{1}_{\{4y \ln(1+y^{-1})<u^{-1}\}} \left(2y + O(y^2) \right) \ln(1 + y^{-1}) \, dy$$

$$= \frac{12}{\pi^2} \left(y^2 \ln(1 + y^{-1}) + O(y^2) \right) \Big|_{0_+}^{4y \ln(1+y^{-1})=u^{-1}} = \frac{3}{4\pi^2} \frac{1}{u^2 \ln u} (1 + o(1)). \qquad (6.33)$$

The second term is bounded above by

$$\frac{12}{\pi^2} \int_0^\infty \mathbb{1}_{\{\frac{2f(y)}{1+2y} <u^{-1}< \frac{2f(y)}{1+y}\}} \left(\frac{2}{1 + y} - \frac{2}{1 + 2y} \right) \ln(1 + y^{-1}) \, dy$$

$$= \frac{12}{\pi^2} \left(y^2 \ln(1 + y^{-1}) + O(y^2) \right) \Big|_{\frac{2f(y)}{1+y}=u^{-1}}^{\frac{2f(y)}{1+2y}=u^{-1}} = o\left(\frac{1}{u^2 \ln u} \right). \qquad (6.34)$$

This proves (6.27). $\square$

Proposition 6.6. *Let $d \geq 3$. There are constants $c_2 > c_1 > 0$, such that for $u \geq 1$,*

$$c_1 u^{-(1+\frac{d}{2})} \leq \mathbf{P}(\mu_j > u) \leq c_2 u^{-(1+\frac{d}{2})}. \qquad (6.35)$$

Proof. The upper bounds in (6.16) for K_1, and the upper and lower bounds for K_0 in Lemma 6.2 imply

$$\int_{K_1 > u K_0} K_0 \, d\boldsymbol{w} \, dz \ll \int_{\substack{\log \max(t^{-\frac{d-2}{2}}, t^{-\frac{d-1}{2}} \varphi) \gg u \\ 0 \leq \varphi \leq \frac{\pi}{2}}} d\boldsymbol{w} \, dz + \int_{\substack{\min(t^{-1}, t^{-1+\frac{1}{d}} (\pi - \varphi)^{-\frac{2}{d}}) \gg u \\ \frac{\pi}{2} \leq \varphi \leq \pi}} d\boldsymbol{w} \, dz.$$

$$(6.36)$$

Using spherical coordinates, we see that the second term is, up to a multiplicative constant, equal to

$$\int_{\substack{0 < z < w < 1, \ \frac{\pi}{2} \leq \varphi \leq \pi \\ \min((1-z)^{-1}, (1-z)^{-1+\frac{1}{d}} (\pi - \varphi)^{-\frac{2}{d}}) \gg u}} w^{d-2} dw \, z^{d-2} dz \, (\sin \varphi)^{d-3} d\varphi. \qquad (6.37)$$

We now substitute $x = 1 - w$, $y = 1 - z$, $\phi = \pi - \varphi$, and integrate x over $(0, y)$, and then over y. This yields (again up to a multiplicative constant)

$$\int_0^{\pi/2} \min\left(u^{-1}, u^{-\frac{d}{d-1}} \phi^{-\frac{2}{d-1}}\right)^2 (\sin \phi)^{d-3} d\phi$$

$$= u^{-2} \int_0^{u^{-1/2}} (\sin \phi)^{d-3} d\phi + u^{-\frac{2d}{d-1}} \int_{u^{-1/2}}^{\pi/2} \phi^{-\frac{4}{d-1}} (\sin \phi)^{d-3} d\phi \qquad (6.38)$$

$$\ll u^{-(1+\frac{d}{2})}.$$

The first term in (6.36) can be dealt with similarly, and yields a lower order contribution. This proves the upper bound

$$\mathbf{P}(\mu_j > u) \ll u^{-(1+\frac{d}{2})}. \qquad (6.39)$$

As to the lower bound in (6.35), we use (6.15) (and the same variable substitutions as above):

$$\int_{K_1 > u K_0} K_0 \, d\boldsymbol{w} \, dz \gg \int_{\substack{\min(t^{-1}, (\pi - \varphi)^{-2}) \gg u \\ \frac{\pi}{2} \leq \varphi \leq \pi}} d\boldsymbol{w} \, dz$$

$$\gg u^{-2} \int_0^{u^{-1/2}} (\sin \phi)^{d-3} d\varphi \qquad (6.40)$$

$$\gg u^{-(1+\frac{d}{2})}.$$

$$\square$$

Proposition 6.7.

$$\mathbf{P}(m_j > u) = \begin{cases} O\left(u^{-2}(\log u)^{-1}\right) \mathbb{1}_{\{u \leq r_j\}} & (d = 2) \\ O\left(u^{-(1+\frac{d}{2})}\right) \mathbb{1}_{\{u \leq r_j\}} & (d \geq 3). \end{cases} \qquad (6.41)$$

Proof. This follows from Propositions 6.5 and 6.6, since $m_j \leq \mu_j$ and $m_j \leq r_j$. $\square$

Proposition 6.8.

$$\mathbf{E}(m_j) = \bar{\xi} + O(r_j^{-1}). \tag{6.42}$$

Proof. We have

$$
\begin{aligned}
\mathbf{E}(m_j) &= \frac{1}{v_{d-1}} \int_{\mathcal{B}_1^{d-1}} \int_{\mathcal{B}_1^{d-1}} K_{1,r_j}(\boldsymbol{w}, \boldsymbol{z}) \, d\boldsymbol{w} \, d\boldsymbol{z} \\
&= \int_0^{r_j} x \Psi_0(x) \, dx \tag{6.43} \\
&= \bar{\xi} - \int_{r_j}^{\infty} x \Psi_0(x) \, dx,
\end{aligned}
$$

and (6.42) follows from the asymptotics (5.11). $\square$

Proposition 6.9.

$$\mathbf{E}(m_j^2) = \begin{cases} O(\log \log j) & (d = 2) \\ O(1) & (d \geq 3). \end{cases} \tag{6.44}$$

Proof. This is a direct corollary of Proposition 6.7.

Proposition 6.10. *For $d \geq 2$ and $j \to \infty$,*

$$\mathbf{E}(a_j^2) = \frac{\Theta_d}{2} \log j + O(\log \log j), \tag{6.45}$$

$$\mathbf{E}(b_j^2) = \frac{\Theta_d}{2} \log j + O(\log \log j). \tag{6.46}$$

Proof. We have

$$\mathbf{E}(b_j^2) = \frac{1}{v_{d-1}} \int_{\mathcal{B}_1^{d-1}} \int_{\mathcal{B}_1^{d-1}} K_{2,r_j}(\boldsymbol{w}, \boldsymbol{z}) \, d\boldsymbol{w} \, d\boldsymbol{z} = \int_0^{r_j} x^2 \Psi_0(x) \, dx, \tag{6.47}$$

and hence (6.46) follows from the asymptotics (5.11). Relation (6.45) is a consequence of Proposition 6.9. $\square$

Proposition 6.11. *For $d \geq 2$,*

$$\mathbf{E}(a_j^4) \leq \mathbf{E}(b_j^4) = O(r_j^2). \tag{6.48}$$

Proof. We have

$$
\begin{aligned}
\mathbf{E}(b_j^4) &= \frac{1}{v_{d-1}} \int_{\mathcal{B}_1^{d-1}} \int_{\mathcal{B}_1^{d-1}} \frac{K_{2,r_j}(\boldsymbol{w}, \boldsymbol{z})^2}{K_0(\boldsymbol{w}, \boldsymbol{z})} \, d\boldsymbol{w} \, d\boldsymbol{z} \\
&\leq 2 \int_0^{r_j} x^4 \Psi_0(x) \, dx,
\end{aligned} \tag{6.49}
$$

and the claim follows from (5.11). $\square$

Proposition 6.12. *Let $d = 2$. For $u \to \infty$,*

$$\mathbf{P}(\alpha_n > u) \sim \frac{1}{2\pi^2 u^2}. \tag{6.50}$$

$$\mathbf{P}(\beta_n > u) \sim \frac{1}{2\pi^2 u^2}. \tag{6.51}$$

Proof. We follow the same strategy as in the proof of Proposition 6.5. For

$$f(y) = \frac{6y^2(1+y)^2}{1+2y} \ln(1 + y^{-1}), \tag{6.52}$$

the variable substitution (6.31) yields

$$
\begin{aligned}
\mathbf{P}(\beta_n > u) &= \frac{12}{\pi^2} \int_0^\infty \int_{\frac{2}{1+2y}}^{\frac{2}{1+y}} \mathbb{1}_{\{x^2 f(y) < u^{-2}\}} \ln(1 + y^{-1}) \, dx \, dy \\
&= \frac{12}{\pi^2} \int_0^\infty \mathbb{1}_{\{\frac{4f(y)}{(1+y)^2} < u^{-2}\}} \left(\frac{2}{1+y} - \frac{2}{1+2y} \right) \ln(1 + y^{-1}) \, dy \\
&\quad + \frac{12}{\pi^2} \int_0^\infty \mathbb{1}_{\{\frac{4f(y)}{(1+2y)^2} < u^{-2} < \frac{4f(y)}{(1+y)^2}\}} \left(\frac{1}{uf(y)} - \frac{2}{1+2y} \right) \ln(1 + y^{-1}) \, dy.
\end{aligned} \tag{6.53}
$$

The leading order contribution comes from the first term, which evaluates to

$$
\begin{aligned}
&\frac{12}{\pi^2} \int_0^\infty \mathbb{1}_{\{\frac{24y^2}{1+2y} \ln(1+y^{-1}) < u^{-2}\}} \left(2y + O(y^2) \right) \ln(1 + y^{-1}) \, dy \\
&= \frac{12}{\pi^2} \left(y^2 \ln(1 + y^{-1}) + O(y^2) \right) \Big|_{0_+}^{\frac{24y^2}{1+2y} \ln(1+y^{-1}) = u^{-2}} = \frac{1}{2\pi^2 u^2} (1 + o(1)).
\end{aligned} \tag{6.54}
$$

This proves (6.51). To see that (6.50) has the same asymptotics, recall that $\beta_n^2 - \alpha_n^2 = \mu_n^2$ and (6.27). $\square$

Proposition 6.13. *Let $d \geq 3$. There are constants $c_2 > c_1 > 0$, such that for $u \geq 1$,*

$$c_1 u^{-2} \leq \mathbf{P}(\alpha_n > u) \leq \mathbf{P}(\beta_n > u) \leq c_2 u^{-2}. \tag{6.55}$$

Proof. We exploit the bounds in Lemma 6.4. For the upper bound,

$$
\int_{K_2 > u^2 K_0} K_0 \, d\mathbf{w} \, dz \ll \int_{\substack{\max(t^{-\frac{d-2}{2}}, t^{-\frac{d-1}{2}} \varphi) \gg u^2 \\ 0 \leq \varphi \leq \frac{\pi}{2}}} d\mathbf{w} \, dz
$$

$$
+ \int_{\substack{\min(t^{-(1+\frac{d}{2})}, t^{-\frac{d+1}{2} + \frac{1}{d}} (\pi - \varphi)^{-(1+\frac{2}{d})}) \gg u^2 \\ \frac{\pi}{2} \leq \varphi \leq \pi}} d\mathbf{w} \, dz. \tag{6.56}
$$

Set

$$\alpha := \frac{4}{d+2}, \qquad \beta := \frac{4d}{d(d+1) - 2}, \qquad \gamma := \frac{2d+4}{d(d+1) - 2}, \tag{6.57}$$

and note that

$$\frac{\alpha - \beta}{\gamma} = -\frac{2}{d+2}. \tag{6.58}$$

Using polar co-ordinates as before, the second term in (6.56) evaluates to

$$\int_0^{\pi/2} \min\left(u^{-\alpha}, u^{-\beta}\phi^{-\gamma}\right)^2 (\sin\phi)^{d-3} d\phi$$

$$= u^{-2\alpha} \int_0^{u^{\frac{\alpha-\beta}{\gamma}}} (\sin\phi)^{d-3} d\phi + u^{-2\beta} \int_{u^{\frac{\alpha-\beta}{\gamma}}}^{\pi/2} \phi^{-2\gamma} (\sin\phi)^{d-3} d\phi \tag{6.59}$$

$$\ll u^{-2\alpha} u^{(d-2)\frac{\alpha-\beta}{\gamma}} + u^{-2\beta} u^{(d-2-2\gamma)\frac{\alpha-\beta}{\gamma}} = 2u^{-2}.$$

A similar calculation shows that the first term in (6.56) produces a lower order contribution. This establishes the upper bound in (6.55). For the lower bound for $\mathbf{P}(\beta_n > u)$,

$$\int_{K_2 > u^2 K_0} K_0 \, d\mathbf{w} \, dz \gg \int_{\substack{\min(t^{-(1+\frac{d}{2})}, (\pi-\varphi)^{-(d+2)}) \gg u^2 \\ \frac{\pi}{2} \leq \varphi \leq \pi}} d\mathbf{w} \, dz$$

$$\gg u^{-2\alpha} \int_0^{u^{-\frac{2}{d+2}}} (\sin\phi)^{d-3} d\phi \tag{6.60}$$

$$\gg u^{-2\alpha} u^{-2\frac{d-2}{d+2}} = u^{-2}.$$

The lower bound for $\mathbf{P}(\alpha_n > u)$ follows by combining the lower bound for $\mathbf{P}(\beta_n > u)$ with (6.35). $\quad\square$

7. Spectral Gaps

Let V be a finite-dimensional real vector space with inner product $\langle \cdot, \cdot \rangle_V$ and norm $\|x\| := \langle x, x \rangle^{1/2}$. Denote by $\mathcal{H} = \mathrm{L}^2(\mathcal{B}_1^{d-1}, V, v_{d-1}^{-1} d\mathbf{w})$ the Hilbert space of square-integrable functions $\mathcal{B}_1^{d-1} \to V$ with inner product

$$\langle f, g \rangle := \frac{1}{v_{d-1}} \int_{\mathcal{B}_1^{d-1}} \langle f(\mathbf{w}), g(\mathbf{w}) \rangle_V \, d\mathbf{w}, \tag{7.1}$$

and norm $\|f\| := \langle f, f \rangle^{1/2}$. We will also denote by $\|\cdot\|$ the corresponding operator norm on $\mathcal{H} \to \mathcal{H}$. In the following, (ρ, V) will denote a representation of $\mathrm{SO}(d)$ with group homomorphism $\rho : \mathrm{SO}(d) \to \mathrm{O}(V)$.

Define the following operators on $\mathcal{H}$:

$$Pf(\mathbf{w}) := \int_{\mathcal{B}_1^{d-1}} K_0(\mathbf{w}, z) f(z) \, dz, \tag{7.2}$$

$$\Pi f(\mathbf{w}) := \frac{1}{v_{d-1}} \int_{\mathcal{B}_1^{d-1}} f(z) \, dz, \tag{7.3}$$

$$Uf(\mathbf{w}) := \rho(S(\mathbf{w})) f(\mathbf{w}). \tag{7.4}$$

We have

$$\Pi P = P \Pi = \Pi. \tag{7.5}$$

Denote by $\mathcal{H}_0 := \Pi \mathcal{H}$ the subspace of constant functions, and by $\mathcal{H}_1 = (I - \Pi)\mathcal{H}$ its orthogonal complement. (This means that all components of $f \in \mathcal{H}_1$ have zero mean.) Note that for $f \in \mathcal{H}_0$ we have $Pf = f$, and for $f \in \mathcal{H}_1$ we have $Pf \in \mathcal{H}_1$.

Proposition 7.1. *The operator P has the spectral gap $1 - \omega_0$ with*

$$\omega_0 := \|P - \Pi\| \leq 1 - \frac{1}{2^d \zeta(d)}. \tag{7.6}$$

Proof. This follows from the standard Doeblin argument. Note that, since $K_0(\boldsymbol{w}, \boldsymbol{z})$ is the kernel of a stochastic transition operator with respect to $d\boldsymbol{w}$ on $\mathcal{B}_1^{d-1}$, we have

$$J := v_{d-1} \inf K_0(\boldsymbol{w}, \boldsymbol{z}) \leq 1. \tag{7.7}$$

If $J = 1$, we have $P = \Pi$ and thus $\omega_0 = 0$. Assume therefore $0 \leq J < 1$. Then

$$Q := (1 - J)^{-1}(P - J\Pi) \tag{7.8}$$

is itself a stochastic transition operator, with the same stationary measure. Using (7.5), we can write

$$P = \Pi + (1 - J)(I - \Pi)Q(I - \Pi), \tag{7.9}$$

and so

$$\|P - \Pi\| \leq (1 - J)\|Q\| = 1 - J. \tag{7.10}$$

The claim of the proposition now follows from (6.10). $\quad\square$

Lemma 7.2. *Let $\theta : [0, 1) \to \mathbb{R}$ be measurable, so that*

$$\text{meas}\{w \in [0, 1) : \theta(w) \notin \pi\mathbb{Q}\} > 0, \tag{7.11}$$

and let (ρ, V) be a non-trivial irreducible representation of $\mathrm{SO}(d)$. Then

$$\delta_\rho := \left\| \frac{1}{v_{d-1}} \int_{\mathcal{B}_1^{d-1}} \rho(S(\boldsymbol{w})) \, d\boldsymbol{w} \right\| < 1. \tag{7.12}$$

Proof. For any fixed $\boldsymbol{e} \in \mathrm{S}_1^{d-1}$, let $\Lambda_{\boldsymbol{e}}$ be the push-forward of Lebesgue measure on $[0, 1)$ under the map

$$[0, 1) \to \mathrm{SO}(d), \qquad w \mapsto S(w\boldsymbol{e}). \tag{7.13}$$

The group generated by the support of $\Lambda_{\boldsymbol{e}}$ is, by assumption (7.11), dense in the subgroup

$$\left\{ E(\phi\boldsymbol{e}) : \phi \in [0, 2\pi) \right\} \simeq \mathrm{SO}(2), \tag{7.14}$$

with $E(\boldsymbol{x})$ as in (2.8). Next, let Λ be the push-forward of Lebesgue measure on $\mathcal{B}_1^{d-1}$ under the map

$$\mathcal{B}_1^{d-1} \to \mathrm{SO}(d), \quad \boldsymbol{w} \mapsto S(\boldsymbol{w}). \tag{7.15}$$

The above observation, together with the fact that

$$\left\{ E(\boldsymbol{x}) : \boldsymbol{x} \in \mathcal{B}_{2\pi}^{d-1} \right\} \tag{7.16}$$

generates $\mathrm{SO}(d)$, implies that the group generated by the support of Λ is dense in $\mathrm{SO}(d)$. The claim now follows from well known arguments [27]. $\quad\square$

Proposition 7.3. *Let θ and (ρ, V) be as in Lemma 7.2. Then the operator PU has spectral radius*

$$\omega_\rho := \lim_{n \to \infty} \|(PU)^n\|^{1/n} < 1. \tag{7.17}$$

Proof. It is sufficient to prove $\|PUP\| < 1$, that is

$$\sup_{f \in \mathcal{H}: f \neq 0} \frac{\|PUPf\|^2}{\|f\|^2} < 1. \tag{7.18}$$

We may restrict to functions of the form $f = \alpha f_0 + f_1$, where $\alpha > 0$ and $f_0 \in \mathcal{H}_0$, $f_1 \in \mathcal{H}_1$ with $\|f_0\| = \|f_1\| = 1$. Note that in this case $\|f\|^2 = \alpha^2 + 1$, and hence the supremum (7.18) equals

$$\sup_{\alpha > 0} \frac{1}{\alpha^2 + 1} \sup_{f_0, f_1} \|PUP(\alpha f_0 + f_1)\|^2. \tag{7.19}$$

Now,

$$\begin{aligned}
\sup_{f_1} \|PUP(\alpha f_0 + f_1)\|^2 &= \sup_{f_1} \|PU(\alpha f_0 + Pf_1)\|^2 \\
&\leq \sup_{f_1} \|PU(\alpha f_0 + \omega_0 f_1)\|^2
\end{aligned} \tag{7.20}$$

since

$$\{Pf_1 : f_1 \in \mathcal{H}_1, \ \|f_1\| = 1\} \subseteq \{f_1 \in \mathcal{H}_1, \ \|f_1\| \leq \omega_0\}. \tag{7.21}$$

We have

$$U(\alpha f_0 + \omega_0 f_1) = \alpha y_0 + \omega_0 y_1 + \alpha \tilde{y}_0 + \omega_0 \tilde{y}_1, \tag{7.22}$$

where

$$y_0 := \Pi U f_0 \in \mathcal{H}_0, \quad \tilde{y}_0 := (I - \Pi)U f_0 \in \mathcal{H}_1, \tag{7.23}$$
$$y_1 := \Pi U f_1 \in \mathcal{H}_0, \quad \tilde{y}_1 := (I - \Pi)U f_1 \in \mathcal{H}_1. \tag{7.24}$$

Therefore,

$$\begin{aligned}
&\|PU(\alpha f_0 + \omega_0 f_1)\|^2 \\
&= \|\alpha y_0 + \omega_0 y_1\|^2 + \|P(\alpha \tilde{y}_0 + \omega_0 \tilde{y}_1)\|^2 \\
&\leq \|\alpha y_0 + \omega_0 y_1\|^2 + \omega_0^2 \|\alpha \tilde{y}_0 + \omega_0 \tilde{y}_1\|^2 \\
&= \alpha^2(\omega_0^2 + (1 - \omega_0^2)\|y_0\|^2) + 2\alpha\omega_0(1 - \omega_0^2)\langle y_0, y_1 \rangle + \omega_0^2(\omega_0^2 + (1 - \omega_0^2)\|y_1\|^2).
\end{aligned} \tag{7.25}$$

In the last equality we have used the relations

$$\langle \tilde{y}_0, \tilde{y}_1 \rangle = -\langle y_0, y_1 \rangle, \tag{7.26}$$

which follows from $\langle U f_0, U f_1 \rangle = \langle f_0, f_1 \rangle = 0$, and

$$\|\tilde{y}_0\|^2 = 1 - \|y_0\|^2, \quad \|\tilde{y}_1\|^2 = 1 - \|y_1\|^2. \tag{7.27}$$

Since f_0 is a constant function with $\|f_0\| = 1$, we have by Lemma 7.2 $\|y_0\| \leq \delta_\rho < 1$. Furthermore $\|y_1\| \leq 1$. This shows

$$\sup_{f \in \mathcal{H}: f \neq 0} \frac{\|PUPf\|^2}{\|f\|^2} \leq \sup_{\alpha > 0} \sup_{\substack{\|y_0\| \leq \delta_\rho \\ \|y_1\| \leq 1}} B(\alpha), \tag{7.28}$$

with

$$B(\alpha) := \frac{\alpha^2(\omega_0^2 + (1 - \omega_0^2)\|y_0\|^2) + 2\alpha\omega_0(1 - \omega_0^2)\langle y_0, y_1\rangle + \omega_0^2(\omega_0^2 + (1 - \omega_0^2)\|y_1\|^2)}{\alpha^2 + 1}. \tag{7.29}$$

The final step in the proof of Proposition 7.3 is now to show that

$$\sup_{\alpha > 0} \sup_{\substack{\|y_0\| \leq \delta_\rho \\ \|y_1\| \leq 1}} B(\alpha) < 1. \tag{7.30}$$

To achieve this, first note that

$$\sup_{\substack{\|y_0\| \leq \delta_\rho \\ \|y_1\| \leq 1}} B(0) = \sup_{\|y_1\| \leq 1} \omega_0^2(\omega_0^2 + (1 - \omega_0^2)\|y_1\|^2) \leq \omega_0^2 < 1 \tag{7.31}$$

and

$$\lim_{\alpha \to \infty} \sup_{\substack{\|y_0\| \leq \delta_\rho \\ \|y_1\| \leq 1}} B(\alpha) = \sup_{\|y_0\| \leq \delta_\rho} \omega_0^2 + (1 - \omega_0^2)\|y_0\|^2 < 1. \tag{7.32}$$

To prove, (7.30), it is therfore sufficient that the quadratic equation

$$B(\alpha) = 1 \tag{7.33}$$

has no positive real solution for all $\|y_0\| \leq \delta_\rho$, $\|y_1\| \leq 1$. This in turn holds, if the discriminant of Eq. (7.33) is strictly negative, i.e.

$$\sup_{\substack{\|y_0\| \leq \delta_\rho \\ \|y_1\| \leq 1}} \left[-4(1 - \omega_0^2)^2 \left\{ 1 - \|y_0\|^2 + \omega_0^2 \left[(1 - \|y_0\|^2)(1 - \|y_1\|^2) - \langle y_0, y_1\rangle^2 \right] \right\} \right] < 0. \tag{7.34}$$

Because $1 - \omega_0^2 > 0$ and $1 - \|y_0\|^2 \geq 1 - \delta_\rho^2 > 0$, it remains to be shown that

$$\langle y_0, y_1\rangle^2 \leq (1 - \|y_0\|^2)(1 - \|y_1\|^2). \tag{7.35}$$

To this end, apply Eqs. (7.26), (7.27) and the Cauchy–Schwarz inequality,

$$|\langle y_0, y_1\rangle| = |\langle \tilde{y}_0, \tilde{y}_1\rangle| \leq \|\tilde{y}_0\|\|\tilde{y}_1\| = \sqrt{(1 - \|y_0\|^2)}\sqrt{(1 - \|y_1\|^2)}, \tag{7.36}$$

which completes the proof. □

8. Exponential Mixing

We will now apply the spectral estimates of the previous section to obtain exponential mixing rates.

Denote by $(\rho_1, V_1) = (\mathrm{id}, \mathbb{R}^d)$ the natural representation of $SO(d)$, and by (ρ_2, V_2) the adjoint representation of $SO(d)$ on the vector space V_2 of real symmetric traceless $d \times d$ matrices defined by

$$\rho_2(R) : M \mapsto RM\,{}^tR. \tag{8.1}$$

The inner product on V_1 is the standard Euclidean inner product

$$\langle x_1, x_2 \rangle_{V_1} := x_1 \cdot x_2, \tag{8.2}$$

and on V_2 the Hilbert–Schmidt inner product

$$\langle M_1, M_2 \rangle_{V_2} := \mathrm{tr}(M_1 M_2). \tag{8.3}$$

Proposition 8.1. *Fix any* $\omega \in [\omega_0, 1) \cap (\omega_{\rho_1}, 1) \cap (\omega_{\rho_2}, 1)$, $m \in \mathbb{N}$ *and* $p = 0, 1, 2$. *Then there is a constant* $C_m > 0$ *such that, for all* $n_1, n_2 \in \mathbb{N}$, $v_0, e \in S_1^{d-1}$ *and all measurable* $f, g : (\mathcal{B}_1^{d-1})^{m+1} \to \mathbb{R}$ *with*

$$\mathbf{E}\left(f(\eta_0, \ldots, \eta_m)^2\right) < \infty, \quad \mathbf{E}\left(g(\eta_0, \ldots, \eta_m)^2\right) < \infty, \tag{8.4}$$

we have

$$\left| \mathbf{Cov}\left((e \cdot V_{n_1})^p f(\eta_{n_1}, \ldots, \eta_{n_1+m}), (e \cdot V_{n_2})^p g(\eta_{n_2}, \ldots, \eta_{n_2+m})\right)\right|$$
$$\leq C_m \,\omega^{|n_1 - n_2| - m} \sqrt{\mathbf{E}\left(f(\eta_0, \ldots, \eta_m)^2\right)} \sqrt{\mathbf{E}\left(g(\eta_0, \ldots, \eta_m)^2\right)}. \tag{8.5}$$

(Note that we have fixed $V_0 = v_0$, which breaks stationarity.)

Proof. Assume without loss of generality that $n_1 \leq n_2$. We have

$$\mathbf{Cov}\left((e \cdot V_{n_1})^p f(\eta_{n_1}, \ldots, \eta_{n_1+m}), (e \cdot V_{n_2})^p g(\eta_{n_2}, \ldots, \eta_{n_2+m})\right)$$
$$= \mathbf{E}\left[\mathbf{Cov}\left((e \cdot V_{n_1})^p f(\eta_{n_1}, \ldots, \eta_{n_1+m}), (e \cdot V_{n_2})^p g(\eta_{n_2}, \ldots, \eta_{n_2+m}) \mid V_{n_1}\right)\right]. \tag{8.6}$$

It is therefore sufficient to prove (8.5) conditioned on V_{n_1} with a constant C_m independent on V_{n_1}, or equivalently, to show that form any $n \in \mathbb{Z}_{\geq 0}$, $v_0, e \in S_1^{d-1}$,

$$\left|\mathbf{Cov}\left(f(\eta_0, \ldots, \eta_m), (e \cdot V_n)^p g(\eta_n, \ldots, \eta_{n+m})\right)\right|$$
$$\leq C_m \,\omega^{n-m} \sqrt{\mathbf{E}\left(f(\eta_0, \ldots, \eta_m)^2\right)} \sqrt{\mathbf{E}\left(g(\eta_0, \ldots, \eta_m)^2\right)}. \tag{8.7}$$

The case $n \leq m$ follows from the Cauchy–Schwarz inequality. We assume therefore in the following that $n > m$.

Case A $p = 0$. Define the functions $\mathcal{B}_1^{d-1} \to \mathbb{R}$

$$\tilde{f}(w) := \mathbf{E}\left(f(\eta_0, \ldots, \eta_m) \mid \eta_m = w\right), \tag{8.8}$$
$$\tilde{g}(w) := \mathbf{E}\left(g(\eta_0, \ldots, \eta_m) \mid \eta_0 = w\right). \tag{8.9}$$

Then

$$\mathbf{Cov}\left(f(\eta_0, \ldots, \eta_m), g(\eta_n, \ldots, \eta_{n+m})\right) = \mathbf{Cov}\left(\tilde{f}(\eta_m), \tilde{g}(\eta_n)\right)$$

$$= \langle (I - \Pi)\tilde{f}, (P^{n-m} - \Pi)\tilde{g} \rangle \quad (8.10)$$

$$= \langle (I - \Pi)\tilde{f}, (P - \Pi)^{n-m}\tilde{g} \rangle,$$

and so

$$\left|\mathbf{Cov}\left(f(\eta_0, \ldots, \eta_m), g(\eta_n, \ldots, \eta_{n+m})\right)\right| \leq \|\tilde{f}\| \|(P - \Pi)^{n-m}\tilde{g}\|$$

$$\leq \omega_0^{n-m} \|\tilde{f}\| \|\tilde{g}\| \quad (8.11)$$

in view of Proposition 7.1. Finally,

$$\|\tilde{f}\|^2 \leq \mathbf{E}\left(f(\eta_0, \ldots, \eta_m)^2\right), \quad (8.12)$$

since $\tilde{f}$ is obtained from f via orthogonal projection, thus decreasing the L^2-norm (likewise for $\tilde{g}$). This proves (8.7) for $p = 0$.

Case B $p = 1$. Set $\tilde{e} = R(v_0)^{-1}e$, and

$$\overline{f} := \mathbf{E}\, f(\eta_0, \ldots, \eta_m). \quad (8.13)$$

Now

$$\mathbf{Cov}\left(f(\eta_0, \ldots, \eta_m), (e \cdot V_n)g(\eta_n, \ldots, \eta_{n+m})\right)$$

$$= \mathbf{Cov}\left(f(\eta_0, \ldots, \eta_m), (\tilde{e} \cdot S(\eta_1) \cdots S(\eta_n)e_1)g(\eta_n, \ldots, \eta_{n+m})\right) \quad (8.14)$$

$$= \mathbf{E}\left([f(\eta_0, \ldots, \eta_m) - \overline{f}](\tilde{e} \cdot S(\eta_1) \cdots S(\eta_n)e_1)g(\eta_n, \ldots, \eta_{n+m})\right).$$

By using the vector-valued functions $\mathcal{B}_1^{d-1} \to \mathbb{R}^d$

$$\tilde{f}(w) := \mathbf{E}\left([f(\eta_0, \ldots, \eta_m) - \overline{f}]\, {}^{\mathrm{t}}S(\eta_m) \cdots {}^{\mathrm{t}}S(\eta_1) \mid \eta_m = w\right)\tilde{e}, \quad (8.15)$$

$$\tilde{g}(w) := \mathbf{E}\left(g(\eta_0, \ldots, \eta_m) \mid \eta_0 = w\right)e_1, \quad (8.16)$$

we find

$$\mathbf{Cov}\left(f(\eta_0, \ldots, \eta_m), (e \cdot V_n)g(\eta_n, \ldots, \eta_{n+m})\right)$$

$$= \mathbf{E}\left(\tilde{f}(\eta_m) \cdot S(\eta_{m+1}) \cdots S(\eta_n)\tilde{g}(\eta_n)\right) \quad (8.17)$$

$$= \langle \tilde{f}, (PU)^{n-m}\tilde{g} \rangle.$$

We conclude from Proposition 7.3 applied to the natural representation (ρ_1, V_1):

$$\left|\mathbf{Cov}\left(f(\eta_0, \ldots, \eta_m), (e \cdot V_n)g(\eta_n, \ldots, \eta_{n+m})\right)\right| \leq c\,\omega^{n-m} \|\tilde{f}\| \|\tilde{g}\| \quad (8.18)$$

for some $c > 0$. Finally, because orthogonal projection decreases the L^2-norm,

$$\|\tilde{f}\|^2 \leq \mathbf{E}\left(\|[f(\eta_0, \ldots, \eta_m) - \overline{f}]\, {}^{\mathrm{t}}S(\eta_m) \cdots {}^{\mathrm{t}}S(\eta_1)\tilde{e}\|_{V_1}^2\right)$$

$$= \mathbf{E}\left([f(\eta_0, \ldots, \eta_m) - \overline{f}]^2\right) \quad (8.19)$$

$$\leq \mathbf{E}\left(f(\eta_0, \ldots, \eta_m)^2\right).$$

Case C $p = 2$. The vector $\tilde{e}$ and the expectation $\overline{f}$ are defined as in Case B. We have

$$
\begin{aligned}
&\mathbf{Cov}\left(f(\eta_0, \ldots, \eta_m), (e \cdot V_n)^2 g(\eta_n, \ldots, \eta_{n+m})\right) \\
&= \mathbf{Cov}\left(f(\eta_0, \ldots, \eta_m), (\tilde{e} \cdot S(\eta_1) \cdots S(\eta_n)e_1)^2 g(\eta_n, \ldots, \eta_{n+m})\right) \quad (8.20) \\
&= \mathbf{E}\left([f(\eta_0, \ldots, \eta_m) - \overline{f}](\tilde{e} \cdot S(\eta_1) \cdots S(\eta_n)e_1)^2 g(\eta_n, \ldots, \eta_{n+m})\right).
\end{aligned}
$$

We write

$$
\begin{aligned}
(\tilde{e} \cdot S(\eta_1) \cdots S(\eta_n)e_1)^2 &= \frac{1}{d} + \mathrm{tr}\left[\tilde{E} S(\eta_1) \cdots S(\eta_n) E_1 \,{}^{\mathrm{t}}(S(\eta_1) \cdots S(\eta_n))\right] \\
&= \frac{1}{d} + \langle \tilde{E}, \rho(S(\eta_1) \cdots S(\eta_n)) E_1 \rangle_{V_2},
\end{aligned} \quad (8.21)
$$

where V_2 is the vector space of symmetric traceless $d \times d$ matrices, and

$$
E_1 := e_1 \otimes e_1 - \frac{1}{d} I_d \in V_2, \quad \tilde{E} := \tilde{e} \otimes \tilde{e} - \frac{1}{d} I_d \in V_2. \quad (8.22)
$$

The constant term $\frac{1}{d}$ in (8.21) contributes to (8.20) the term

$$
\begin{aligned}
&\frac{1}{d}\left|\mathbf{E}\left([f(\eta_0, \ldots, \eta_m) - \overline{f}]g(\eta_n, \ldots, \eta_{n+m})\right)\right| \\
&\qquad \le \frac{1}{d}\, \omega_0^{n-m} \sqrt{\mathbf{E}\left(f(\eta_0, \ldots, \eta_m)^2\right)}\sqrt{\mathbf{E}\left(g(\eta_0, \ldots, \eta_m)^2\right)}. \quad (8.23)
\end{aligned}
$$

This follows from our discussion in Case A ($p = 0$). The non-constant term in (8.21) is handled in analogy with Case B. Define functions $\mathcal{B}_1^{d-1} \to V_2$

$$
\tilde{f}(\boldsymbol{w}) := \mathbf{E}\left([f(\eta_0, \ldots, \eta_m) - \overline{f}]\rho({}^{\mathrm{t}}S(\eta_m) \cdots {}^{\mathrm{t}}S(\eta_1)) \mid \eta_m = \boldsymbol{w}\right) \tilde{E}, \quad (8.24)
$$

$$
\tilde{g}(\boldsymbol{w}) := \mathbf{E}\left(g(\eta_0, \ldots, \eta_m) \mid \eta_0 = \boldsymbol{w}\right) E_1, \quad (8.25)
$$

so that the non-constant contribution to (8.20) becomes

$$
\begin{aligned}
&\mathbf{E}\left([f(\eta_0, \ldots, \eta_m) - \overline{f}](\tilde{e} \cdot S(\eta_1) \cdots S(\eta_n)e_1)^2 g(\eta_n, \ldots, \eta_{n+m})\right) \\
&= \mathbf{E}\left([f(\eta_0, \ldots, \eta_m) - \overline{f}]\langle \tilde{E}, \rho(S(\eta_1) \cdots S(\eta_n)) E_1 \rangle_{V_2} g(\eta_n, \ldots, \eta_{n+m})\right)(8.26) \\
&= \langle \tilde{f}, (PU)^{n-m} \tilde{g} \rangle.
\end{aligned}
$$

We now apply Proposition 7.3 with the adjoint representation (ρ_2, V_2). This yields the desired bound. $\square$

We will also require the following estimate.

Proposition 8.2. *Fix ω and m as in Proposition 8.1. Then there is a constant $\widetilde{C} > 0$ such that, for all $n \in \mathbb{N}$, $\boldsymbol{v}_0, e \in \mathrm{S}_1^{d-1}$ and all measurable $g : (\mathcal{B}_1^{d-1})^{m+1} \to \mathbb{R}$ with*

$$
\mathbf{E}\left(g(\eta_0, \ldots, \eta_m)^2\right) < \infty, \quad (8.27)
$$

we have

$$
\left|\mathbf{E}\left((e \cdot V_n)g(\eta_n, \ldots, \eta_{n+m})\right)\right| \le \widetilde{C}\, \omega^n \sqrt{\mathbf{E}\left(g(\eta_0, \ldots, \eta_m)^2\right)}. \quad (8.28)
$$

$$
\left|\mathbf{E}\left([(e \cdot V_n)^2 - d^{-1}]g(\eta_n, \ldots, \eta_{n+m})\right)\right| \le \widetilde{C}\, \omega^n \sqrt{\mathbf{E}\left(g(\eta_0, \ldots, \eta_m)^2\right)}. \quad (8.29)
$$

Proof. With $\tilde{g}$ defined as in the previous proof, Case B, we have

$$\mathbf{E}\left((e \cdot V_n)g(\eta_n, \dots, \eta_{n+m})\right) = \langle 1, (PU)^n \tilde{g}\rangle. \tag{8.30}$$

The bound (8.28) now follows from Proposition 7.3. Relation (8.29) follows similarly from Case C of the previous proof. □

9. Proof of the Main Lemmas

We now turn to the proofs of the four main lemmas in Sect. 4.

Proof of Lemma 4.1. We have

$$\|Q_n - Q_n'\| \leq \sum_{j=1}^{n} \zeta_j \tag{9.1}$$

with

$$\zeta_j := \xi_j \mathbb{1}_{\{\xi_j^2 > j(\log j)^\gamma\}}. \tag{9.2}$$

By (5.11), we have

$$\mathbf{P}(\zeta_j \neq 0) = O\left(j^{-1}(\log j)^{-\gamma}\right) \tag{9.3}$$

This is summable (since $\gamma > 1$) and so, by the Borel–Cantelli lemma, $\zeta_j \neq 0$ only for finitely many j. This proves Lemma 4.1. □

Proof of Lemma 4.2. Set $\zeta_j := (e \cdot V_{j-1}) \, m_j$. We need to show that, for every $e \in \mathrm{S}_1^{d-1}$, the sequence of random variables

$$\begin{cases} \dfrac{\sum_{j=1}^n \zeta_j}{\sqrt{n \log \log n}} & (d = 2) \\[2ex] \dfrac{\sum_{j=1}^n \zeta_j}{\sqrt{n}} & (d \geq 3) \end{cases} \tag{9.4}$$

is tight. Now choose in Proposition 8.2 $m = 1$ and $g(w, z) = K_{1,r_j}(w, z)/K_0(w, z)$, and use Proposition 6.9 to bound $\mathbf{E}(g(\eta_0, \eta_1)^2) = \mathbf{E}(m_j^2) = O(\log \log j)$ for $d = 2$, and $= O(1)$ for $d \geq 3$. This proves $\mathbf{E}(\zeta_j) = O(\omega^j)$. Therefore

$$\mathbf{Var}(\zeta_j^2) = \mathbf{E}(\zeta_j^2) + O(\omega^{2j}). \tag{9.5}$$

Proposition 6.9 yields

$$\mathbf{E}(\zeta_j^2) \leq \mathbf{E}(m_j^2) = \begin{cases} O(\log \log j) & (d = 2) \\ O(1) & (d \geq 3). \end{cases} \tag{9.6}$$

Due to Proposition 8.1, we also have

$$\mathbf{Cov}(\zeta_i, \zeta_j) \leq \sqrt{\mathbf{E}(\zeta_i^2)}\sqrt{\mathbf{E}(\zeta_j^2)}\,\omega^{|i-j|}. \tag{9.7}$$

Hence

$$\mathbf{E}\Big[\Big(\sum_{j=1}^{n}\zeta_j\Big)^2\Big] = \begin{cases} O(n\log\log n) & (d=2) \\ O(n) & (d\geq 3), \end{cases} \tag{9.8}$$

which establishes the tightness of (9.4) and thus Lemma 4.2. □

Proof of Lemma 4.3. Let $\zeta_j := (e \cdot V_{j-1})^2 a_j^2$. It is sufficient to prove that for any unit vector $e \in S_1^{d-1}$

$$\frac{\sum_{j=1}^{n}\zeta_j}{n\log n} \xrightarrow{\ \mathbf{P}\ } \sigma_d^2. \tag{9.9}$$

Using (8.29) in Proposition 8.2, and Proposition 6.10,

$$\mathbf{E}(\zeta_j) = \frac{\Theta_d}{2d}\log j + O(\log\log j), \tag{9.10}$$

and, by Proposition 6.11,

$$\mathbf{E}(a_j^4) = O(j(\log j)^\gamma). \tag{9.11}$$

Furthermore, due to Proposition 8.1 ($p=2$), we have

$$\mathbf{Cov}(\zeta_i,\zeta_j) \leq \sqrt{\mathbf{E}(a_i^4)}\sqrt{\mathbf{E}(a_j^4)}\,\omega^{|i-j|}. \tag{9.12}$$

Hence,

$$\mathbf{E}\Big(\sum_{j=1}^{n}\zeta_j\Big) = \frac{\Theta_d}{2d}n\log n + O(n\log\log n) \tag{9.13}$$

and

$$\mathbf{Var}\Big(\sum_{j=1}^{n}\zeta_j\Big) = O(n^2(\log n)^\gamma) = o((n\log n)^2) \tag{9.14}$$

since $\gamma < 2$. This proves Lemma 4.3. □

Proof of Lemma 4.4. In view of the asymptotic relation for A_n in (4.19) we have to prove that for any $\varepsilon > 0$

$$\frac{\sum_{j=1}^{n}\mathbf{E}(\tilde{\xi}_j^2\mathbb{1}_{\{\tilde{\xi}_j^2>\varepsilon^2 n\log n\}}\mid\boldsymbol{\eta})}{n\log n} \xrightarrow{\ \mathbf{P}\ } 0. \tag{9.15}$$

The *lower tail* $\tilde{\xi}_j < -\varepsilon\sqrt{n\log n}$ is estimated by

$$\begin{aligned}\mathbf{E}\big((\xi_j'-m_j)^2\mathbb{1}_{\{\xi_j'-m_j<-\varepsilon\sqrt{n\log n}\}}\mid\boldsymbol{\eta}\big) &\leq m_j^2\mathbb{1}_{\{m_j>\varepsilon\sqrt{n\log n}\}}\\ &\leq m_j^2\mathbb{1}_{\{m_j>\varepsilon\sqrt{j\log j}\}}.\end{aligned} \tag{9.16}$$

Proposition 6.7 yields

$$\mathbf{P}\big(m_j > \varepsilon\sqrt{j\log j}\big) = \begin{cases} O((\varepsilon^2 j(\log j)^2)^{-1}) & (d=2) \\ O((\varepsilon^2 j\log j)^{-(\frac{1}{2}+\frac{d}{4})}) & (d\geq 3). \end{cases} \tag{9.17}$$

Since this is summable, we have, by the Borel–Cantelli lemma,

$$\frac{\sum_{j=1}^{n}\mathbf{E}\big(\tilde{\xi}_j^2\mathbb{1}_{\{\tilde{\xi}_j<-\varepsilon\sqrt{n\log n}\}}\,\big|\,\boldsymbol{\eta}\big)}{n\log n}\ \xrightarrow{\text{a.s.}}\ 0. \tag{9.18}$$

For the *upper tail* $\tilde{\xi}_j > \varepsilon\sqrt{n\log n}$ we have

$$\begin{aligned}
\zeta_{n,j} &:= \mathbf{E}\big((\xi_j'-m_j)^2\mathbb{1}_{\{\xi_j-m_j>\varepsilon\sqrt{n\log n}\}}\,\big|\,\boldsymbol{\eta}\big) \\
&\leq \mathbf{E}\big(\xi_j'^2\mathbb{1}_{\{\xi_j'>\varepsilon\sqrt{n\log n}\}}\,\big|\,\boldsymbol{\eta}\big) \\
&= \mathbf{E}\big(\xi_j^2\mathbb{1}_{\{\varepsilon^2 n\log n<\xi_j^2\leq j(\log j)^\gamma\}}\,\big|\,\boldsymbol{\eta}\big).
\end{aligned} \tag{9.19}$$

On the other hand, in view of (5.11), we have for $n\to\infty$,

$$\mathbf{E}\big(\zeta_{n,j}\big) \ll \log\sqrt{\frac{n(\log n)^\gamma}{\varepsilon^2 n\log n}}\ \sim\ \frac{\gamma-1}{2}\log\log n, \tag{9.20}$$

and therefore

$$\frac{\mathbf{E}\big(\sum_{j=1}^{n}\zeta_{n,j}\big)}{n\log n}\ \to\ 0. \tag{9.21}$$

From (9.18) and (9.21), the assertion of Lemma 4.4 follows. □

10. General Initial Data

Up to now we have assumed that $(\xi_1,\boldsymbol{\eta}_1)$ has density $\Psi_0(x,z)$. We now extend the above results to more general initial data $(\xi_1,\boldsymbol{\eta}_1)$, where the only assumption is that the marginal distribution of $\boldsymbol{\eta}_1$ is absolutely continuous with respect to Lebesgue measure on $\mathcal{B}_1^{d-1}$.

Proof of Theorem 3.2 (ii) for general initial data. Since

$$\frac{\xi_1 V_0}{\sqrt{n\log n}}\ \xrightarrow{\mathbf{P}}\ \mathbf{0}, \tag{10.1}$$

it is sufficient to show that

$$\frac{\sum_{j=2}^{n}\xi_j V_{j-1}}{\sigma_d\sqrt{n\log n}}\ \Rightarrow\ \mathcal{N}(0, I_d) \tag{10.2}$$

where $\boldsymbol{\eta}_1$ has (by assumption) an absolutely continuous distribution and $\xi_1 = 0$. By an obvious re-labelling, this is equivalent to showing that

$$\frac{\sum_{j=1}^{n}\xi_j V_{j-1}}{\sigma_d\sqrt{n\log n}}\ \Rightarrow\ \mathcal{N}(0, I_d) \tag{10.3}$$

where η_0 has an absolutely continuous distribution. In view of the remarks following Eq. (4.1), the only difference from the proof of Theorem 3.2 is now that η_0 is distributed according to an absolutely continuous probability measure, rather than Lebesgue measure. Because *tightness*, *almost sure convergence* and *convergence in probability* continue to hold when passing from a measure to a measure which is absolutely continuous with respect to the first, the Lemmas in Sect. 4 remain valid also in the present setting. The proof of Theorem 3.2 for general initial data therefore follows from these lemmas in the same way as for the density $\Psi_0(x, z)$, as described at the end of Sect. 4. □

The following proposition shows that, if (ξ_1, η_1) has density $\Psi(x, z)$ (which appears in the continuous-time setting of the Boltzmann–Grad limit, Theorem 3.1 (i)), then the marginal distribution of η_1 is absolutely continuous. Let

$$\overline{\Psi}(z) := \int_0^\infty \Psi(x, z)\, dx \tag{10.4}$$

Proposition 10.1.

$$\overline{\Psi} \in \mathrm{L}^1(\mathcal{B}_1^{d-1}, dz). \tag{10.5}$$

Proof. The function $\Psi(x, z)$ is continuous, and in view of Corollary 1.2 and Theorem 1.11 in [22], uniformly bounded. The latter theorem produces a precise asymptotics of $\Psi(x, z)$, which implies

$$\Psi(x, z) = O(x^{-2+\frac{2}{d}}), \tag{10.6}$$

uniformly for all $x > 0$, $z \in \mathcal{B}_1^{d-1}$. Thus, for $d \geq 3$, $\overline{\Psi}(z)$ is uniformly bounded, and hence $\overline{\Psi} \in \mathrm{L}^1(\mathcal{B}_1^{d-1}, dz)$ as required.

In dimension $d = 2$, there is an explicit formula for $\Psi(x, z)$, cf. [19, Eq. (30)], which yields (see the last displayed equation of that paper) for $x \to \infty$ and $z \in (-1, 1)$,

$$\Psi(x, z) = \frac{3}{2\pi^2}(1 - u)^2 x^{-1} + O(x^{-2}) \tag{10.7}$$

if $u := x(1 - |z|) \in [0, 1)$, and

$$\Psi(x, z) = 0 \tag{10.8}$$

if $x(1 - |z|) \notin [0, 1)$. The implied constant in (10.7) is independent of x and u. The above asymptotics (and the fact that $\Psi(x, z)$ is uniformly bounded) imply

$$\overline{\Psi}(z) := \int_0^\infty \Psi(x, z)\, dx = \log \frac{1}{1 - |z|} + O(1), \tag{10.9}$$

which holds uniformly for all $z \in (-1, 1)$. We conclude that $\overline{\Psi} \in \mathrm{L}^1((-1, 1), dz)$. □

11. From Discrete to Continuous Time

The following proposition, together with Theorem 3.2 (ii), immediately implies Theorem 3.2 (i). Let us denote by

$$n_t := \left\lfloor \bar{\xi}^{-1} t \right\rfloor \tag{11.1}$$

the (integer part of the) expected number of collisions within time t.

Proposition 11.1. *For any $\varepsilon > 0$*

$$\frac{\|X_t - \mathbf{Q}_{n_t}\|}{t^{5/12+\varepsilon}} \xrightarrow{\mathbf{P}} 0, \tag{11.2}$$

as $t \to \infty$.

By the same argument as in the proof of Theorem 3.2 (see Section 10), it is sufficient to prove Proposition 11.1 in the case when $(\xi_1, \boldsymbol{\eta}_1)$ has density $\Psi_0(x, z)$. We will assume this from now on. Furthermore, note that the left hand side of (11.2) is independent of the choice of $\boldsymbol{v}_0$. We may therefore assume without loss of generality that $\boldsymbol{v}_0$ is a random variable uniformly distributed in S_1^{d-1} (this will only be used in the justification of rel. (11.31) below). The proof of Proposition 11.1, which is given at the end of this section, exploits the following three lemmas.

Lemma 11.2.

$$\frac{\tau_n - n\bar{\xi}}{\sigma_d \sqrt{d\, n \log n}} \Rightarrow \mathcal{N}(0, 1), \tag{11.3}$$

as $n \to \infty$.

Proof. This is a simple variant of the proof of Theorem 3.2 (ii). □

Lemma 11.3. *For all $n \in \mathbb{N}$ and $u \geq 2$,*

$$\mathbf{P}(\|\mathbf{Q}_n\| > u) = O\left(\frac{n \log u}{u^2}\right). \tag{11.4}$$

Proof. We begin by observing that

$$\mathbf{P}(\|\mathbf{Q}_n\| > u) \leq \mathbf{P}\left(\left\|\sum_{j=1}^n \xi_j^* \mathbf{V}_{j-1}\right\| > \frac{u}{2}\right) + \mathbf{P}\left(\left\|\sum_{j=1}^n \mu_j \mathbf{V}_{j-1}\right\| > \frac{u}{2}\right), \tag{11.5}$$

where $\xi_j^* = \xi_j - \mu_j$ and μ_j is the conditional expectation of the (untruncated) ξ_j defined in (6.3). Recall also the definition of the corresponding conditional variance α_j in (6.5). Now,

$$
\begin{aligned}
\mathbf{P}\left(\left\|\sum_{j=1}^n \xi_j^* \mathbf{V}_{j-1}\right\| > \frac{u}{2}\right) &\leq \mathbf{P}\left(\left\|\sum_{j=1}^n \xi_j^* \mathbf{V}_{j-1} \mathbb{1}_{\{\alpha_j \leq u\}}\right\| > \frac{u}{4}\right) \\
&\quad + \mathbf{P}\left(\left\|\sum_{j=1}^n \xi_j^* \mathbf{V}_{j-1} \mathbb{1}_{\{\alpha_j > u\}}\right\| > \frac{u}{4}\right) \\
&\leq \frac{16}{u^2} \sum_{j=1}^n \mathbf{E}\left(\alpha_j^2 \mathbb{1}_{\{\alpha_j \leq u\}}\right) + \sum_{j=1}^n \mathbf{P}(\alpha_j > u) \\
&= O\left(\frac{n \log u}{u^2}\right) + O\left(\frac{n}{u^2}\right),
\end{aligned}
\tag{11.6}
$$

where we have used Chebyshev's inequality and, in the last bound, Propositions 6.12 ($d = 2$) and 6.13 ($d \geq 3$).

The second term in (11.5) is bounded similarly: We have

$$\mathbf{P}\big(\big\| \sum_{j=1}^{n} \mu_j V_{j-1} \big\| > \frac{u}{2}\big) \leq \mathbf{P}\big(\big\| \sum_{j=1}^{n} \mu_j V_{j-1} \mathbb{1}_{\{\mu_j \leq u\}} \big\| > \frac{u}{4}\big)$$

$$+ \mathbf{P}\big(\big\| \sum_{j=1}^{n} \mu_j V_{j-1} \mathbb{1}_{\{\mu_j > u\}} \big\| > \frac{u}{4}\big) \tag{11.7}$$

$$\leq \frac{16}{u^2} \mathbf{E}\big(\big\| \sum_{j=1}^{n} \mu_j V_{j-1} \mathbb{1}_{\{\mu_j \leq u\}} \big\|^2\big) + \sum_{j=1}^{n} \mathbf{P}\big(\mu_j > u\big).$$

To control the first term on the right hand side of (11.7), it is sufficient to bound

$$\mathbf{E}\big[\big(\sum_{j=1}^{n} \zeta_j\big)^2\big], \qquad \zeta_j := (e \cdot V_{j-1}) \mu_j \mathbb{1}_{\{\mu_j \leq u\}}, \tag{11.8}$$

for arbitrary $e \in S_1^{d-1}$. We follow the same steps as in the proof of Lemma 4.2. Let us first show that

$$\mathbf{E}(\zeta_j) = \begin{cases} O(\omega^j \sqrt{\log \log u}) & (d = 2) \\ O(\omega^j) & (d \geq 3). \end{cases} \tag{11.9}$$

To this end choose in Proposition 8.2 $m = 1$ and

$$g(\mathbf{w}, \mathbf{z}) = \frac{K_1(\mathbf{w}, \mathbf{z})}{K_0(\mathbf{w}, \mathbf{z})} \mathbb{1}_{\{K_1(\mathbf{w}, \mathbf{z}) \leq u K_0(\mathbf{w}, \mathbf{z})\}}. \tag{11.10}$$

Propositions 6.5 and 6.6 yield

$$\mathbf{E}(g(\boldsymbol{\eta}_0, \boldsymbol{\eta}_1)^2) = \mathbf{E}(\mu_j^2 \mathbb{1}_{\{\mu_j \leq u\}}) = \begin{cases} O(\log \log u) & (d = 2) \\ O(1) & (d \geq 3), \end{cases} \tag{11.11}$$

and hence (11.9).

The above implies

$$\mathbf{Var}(\zeta_j^2) = \mathbf{E}(\zeta_j^2) + \begin{cases} O(\omega^{2j} \log \log u) & (d = 2) \\ O(\omega^{2j}) & (d \geq 3), \end{cases} \tag{11.12}$$

where

$$\mathbf{E}(\zeta_j^2) \leq \mathbf{E}(\mu_j^2 \mathbb{1}_{\{\mu_j \leq u\}}) = \begin{cases} O(\log \log u) & (d = 2) \\ O(1) & (d \geq 3). \end{cases} \tag{11.13}$$

Due to Proposition 8.1, we also have

$$\mathbf{Cov}(\zeta_i, \zeta_j) \leq \sqrt{\mathbf{E}(\zeta_i^2)}\sqrt{\mathbf{E}(\zeta_j^2)}\omega^{|i-j|}. \tag{11.14}$$

We conclude

$$\mathbf{E}\Big[\Big(\sum_{j=1}^{n}\zeta_j\Big)^2\Big] = \begin{cases} O(n\log\log u) & (d=2) \\ O(n) & (d\geq 3). \end{cases} \tag{11.15}$$

The second term on the right hand side of (11.7) is controlled by the tail estimates in Propositions 6.5 and 6.6. The overall result is

$$\mathbf{P}\Big(\Big\|\sum_{j=1}^{n}\mu_j V_{j-1}\Big\| > \frac{u}{2}\Big) = \begin{cases} O\Big(\dfrac{n\log\log u}{u^2}\Big) + O\Big(\dfrac{n}{u^2\log u}\Big) & (d=2) \\ O\Big(\dfrac{n}{u^2}\Big) + O\Big(\dfrac{n}{u^{1+\frac{d}{2}}}\Big) & (d\geq 3), \end{cases} \tag{11.16}$$

which completes the proof of the lemma. □

Lemma 11.4. *For any $\varepsilon > 0$ and $\delta > 0$,*

$$\lim_{t\to\infty}\mathbf{P}\big(|v_t - n_t| > \delta t^{1/2+\varepsilon}\big) = 0, \tag{11.17}$$

$$\lim_{n\to\infty}\mathbf{P}\big(\max_{1\leq j\leq n}\xi_j > \delta n^{1/2+\varepsilon}\big) = 0, \tag{11.18}$$

$$\lim_{n\to\infty}\mathbf{P}\big(\max_{1\leq m\leq n}\|\mathbf{Q}_m\| > \delta n^{5/6+\varepsilon}\big) = 0. \tag{11.19}$$

Proof of (11.17). Note that, for any $N \in \mathbb{Z}_{\geq 0}, t \geq 0$,

$$v_t \geq N \quad \Leftrightarrow \quad \tau_N \leq t, \tag{11.20}$$

and therefore, with $N(t) := \lfloor n_t + \delta t^{1/2+\varepsilon}\rfloor$,

$$\mathbf{P}\big(v_t - n_t > \delta t^{1/2+\varepsilon}\big) \leq \mathbf{P}\big(v_t \geq N(t)\big) = \mathbf{P}\big(\tau_{N(t)} \leq t\big). \tag{11.21}$$

On the other hand, it follows from Lemma 11.2 that

$$\lim_{t\to\infty}\mathbf{P}\big(\tau_{N(t)} \leq t\big) = 0. \tag{11.22}$$

Similarly, for $M(t) := \lfloor n_t - \delta t^{1/2+\varepsilon}\rfloor$, we have

$$\mathbf{P}\big(v_t - n_t < -\delta t^{1/2+\varepsilon}\big) = \mathbf{P}\big(\tau_{M(t)} > t\big), \tag{11.23}$$

and Lemma 11.2 implies

$$\lim_{t\to\infty}\mathbf{P}\big(\tau_{M(t)} > t\big) = 0. \tag{11.24}$$

□

Proof of (11.18). We use the simplest union bound and Markov's inequality:

$$\mathbf{P}\big(\max_{1\leq j\leq n}\xi_j > \delta n^{1/2+\varepsilon}\big) \leq n\mathbf{P}\big(\xi_1 > \delta n^{1/2+\varepsilon}\big) \leq n\frac{\mathbf{E}\big(\xi_1^{2-\varepsilon}\big)}{\delta^{2-\varepsilon}n^{(1/2+\varepsilon)(2-\varepsilon)}}. \tag{11.25}$$

This sequence converges to 0 for $\varepsilon < \frac{3}{2}$. □

Proof of (11.19). Note first that

$$\max_{1\leq m\leq n} \| \boldsymbol{Q}_m \| \leq \max_{1\leq m\leq n^{2/3}} \| \boldsymbol{Q}_{m\lfloor n^{1/3}\rfloor} \| + n^{1/3} \max_{1\leq m\leq n} \xi_m. \tag{11.26}$$

Hence

$$\mathbf{P}\Big(\max_{1\leq m\leq n} \| \boldsymbol{Q}_m \| > \delta n^{5/6+\varepsilon}\Big)$$

$$\leq \mathbf{P}\Big(\max_{1\leq m\leq n^{2/3}} \| \boldsymbol{Q}_{m\lfloor n^{1/3}\rfloor} \| > \frac{\delta}{2} n^{5/6+\varepsilon}\Big) + \mathbf{P}\Big(\max_{1\leq m\leq n} \xi_m > \frac{\delta}{2} n^{1/2+\varepsilon}\Big). \tag{11.27}$$

The second term on the right hand side of (11.27) converges to zero, due to (11.18). From (11.4) it follows that

$$\mathbf{P}\Big(\max_{1\leq m\leq n^{2/3}} \| \boldsymbol{Q}_{m\lfloor n^{1/3}\rfloor} \| > \delta n^{5/6+\varepsilon}\Big) \ll \frac{\log n}{\delta^2 n^{5/3+2\varepsilon}} \sum_{m=1}^{\lfloor n^{2/3}\rfloor} mn^{1/3} \ll \frac{\log n}{\delta^2 n^{2\varepsilon}} \to 0.$$

$$\tag{11.28}$$

This completes the proof of Lemma 11.4. □

Proof of Proposition 11.1. Since

$$\| X_t - \boldsymbol{Q}_{n_t} \| \leq \| \boldsymbol{Q}_{\nu_t} - \boldsymbol{Q}_{n_t} \| + \xi_{\nu_t+1}, \tag{11.29}$$

we have

$$\mathbf{P}\big(\| X_t - \boldsymbol{Q}_{n_t} \| > \delta t^{5/12+\varepsilon}\big) \leq \mathbf{P}\big(|\nu_t - n_t| > t^{1/2+\varepsilon}\big)$$

$$+ \mathbf{P}\Big(\max_{|m|\leq t^{1/2+\varepsilon}} \xi_{n_t+m} > \frac{\delta}{2} t^{5/12+\varepsilon}\Big) + \mathbf{P}\Big(\max_{|m|\leq t^{1/2+\varepsilon}} \| \boldsymbol{Q}_{n_t+m} - \boldsymbol{Q}_{n_t} \| > \frac{\delta}{2} t^{5/12+\varepsilon}\Big),$$

$$\tag{11.30}$$

and therefore, by stationarity of the Markov process (3.2) (recall that here we may assume without loss of generality that $\boldsymbol{v}_0$ is uniformly distributed in S_1^{d-1}),

$$\mathbf{P}\big(\| X_t - \boldsymbol{Q}_{n_t} \| > \delta t^{5/12+\varepsilon}\big) \leq \mathbf{P}\big(|\nu_t - n_t| > t^{1/2+\varepsilon}\big)$$

$$+ \mathbf{P}\Big(\max_{1\leq m\leq 2t^{1/2+\varepsilon}} \xi_m > \frac{\delta}{2} t^{5/12+\varepsilon}\Big) + \mathbf{P}\Big(\max_{1\leq m\leq 2t^{1/2+\varepsilon}} \| \boldsymbol{Q}_m \| > \frac{\delta}{4} t^{5/12+\varepsilon}\Big). \tag{11.31}$$

The three terms on the right hand side of (11.31) are controlled by Lemma 11.4. This completes the proof of Proposition 11.1. □

12. Convergence of Finite-dimensional Distributions

The convergence of finite-dimensional distribution follows from analogous arguments as in the one-dimensional case (cf. Sect. 4). We include a sketch of the main steps.

We will assume for the rest of this paper that $(\xi_1, \boldsymbol{\eta}_1)$ has density $\Psi_0(x, z)$; by the arguments of Sect. 10 this is without loss of generality.

Proposition 12.1. *Let $d \geq 2$, $\boldsymbol{v}_0 \in S_1^{d-1}$ and assume that the marginal distribution of $\boldsymbol{\eta}_1$ is absolutely continuous. Then, for every fixed k-tuple $(t_1, \ldots, t_k) \in (0, 1]^k$ as $n \to \infty$,*

$$\left(\boldsymbol{Y}_n(t_1), \ldots, \boldsymbol{Y}_n(t_k)\right) \Rightarrow \left(\boldsymbol{W}(t_1), \ldots, \boldsymbol{W}(t_k)\right). \tag{12.1}$$

Proof. We may assume $t_0 := 0 < t_1 < t_2 < \ldots < t_k \leq 1$. The weak convergence (12.1) is equivalent to

$$\begin{aligned}
\left(\boldsymbol{Y}_n(t_1) - \boldsymbol{Y}_n(t_0), \ldots, \boldsymbol{Y}_n(t_k) - \boldsymbol{Y}_n(t_{k-1})\right) \\
\Rightarrow \left(\boldsymbol{W}(t_1) - \boldsymbol{W}(t_0), \ldots, \boldsymbol{W}(t_k) - \boldsymbol{W}(t_{k-1})\right).
\end{aligned} \tag{12.2}$$

Define the kd-dimensional vector

$$\boldsymbol{U}_{j,n} := \begin{pmatrix} \mathbb{1}_{\{j \leq \lfloor t_1 n \rfloor\}} \boldsymbol{V}_j \\ \mathbb{1}_{\{\lfloor t_1 n \rfloor < j \leq \lfloor t_2 n \rfloor\}} \boldsymbol{V}_j \\ \vdots \\ \mathbb{1}_{\{\lfloor t_{k-1} n \rfloor < j \leq \lfloor t_k n \rfloor\}} \boldsymbol{V}_j \end{pmatrix} \tag{12.3}$$

and

$$\boldsymbol{R}_n := \sum_{j=1}^n \xi_j \boldsymbol{U}_j = \begin{pmatrix} \boldsymbol{Q}_n(t_1) - \boldsymbol{Q}_n(t_0) \\ \vdots \\ \boldsymbol{Q}_n(t_k) - \boldsymbol{Q}_n(t_{k-1}) \end{pmatrix}, \quad n \in \mathbb{N}. \tag{12.4}$$

We thus need to show that

$$\frac{\boldsymbol{R}_n}{\sigma_d \sqrt{n \log n}} \Rightarrow \begin{pmatrix} \boldsymbol{W}_n(t_1) - \boldsymbol{W}_n(t_0) \\ \vdots \\ \boldsymbol{W}_n(t_k) - \boldsymbol{W}_n(t_{k-1}) \end{pmatrix}. \tag{12.5}$$

We truncate $\boldsymbol{R}_n$ by defining the random variable

$$\boldsymbol{R}_n' := \sum_{j=1}^n \xi_j' \boldsymbol{U}_{j-1} \tag{12.6}$$

with ξ_j' as in (4.8).

Lemma 12.2. *We have*

$$\sup_{n \in \mathbb{N}} \|\boldsymbol{R}_n - \boldsymbol{R}_n'\| < \infty \tag{12.7}$$

almost surely.

This statement is an immediate consequence of Lemma 4.1, where the bound is established for each component. To prove the central limit theorem for $\boldsymbol{R}_n'$, we consider $\tilde{\xi}_j = \xi_j' - m_j$ with the conditional expectation m_j as in (4.11), and let

$$\widetilde{\boldsymbol{R}}_n := \sum_{j=1}^n \tilde{\xi}_j \boldsymbol{U}_{j-1}. \tag{12.8}$$

Lemma 12.3. *The sequence of random variables*

$$\frac{R'_n - \widetilde{R}_n}{\sqrt{n \log \log n}} \tag{12.9}$$

is tight if $d = 2$, and

$$\frac{R'_n - \widetilde{R}_n}{\sqrt{n}} \tag{12.10}$$

is tight if $d \geq 3$.

This lemma follows directly from Lemma 4.2. Lemmas 12.2 and 12.3 imply that it is sufficient to prove Proposition 12.1 for $\widetilde{R}_n$ in place of R_n. Let us turn to the covariance and recall the definition (4.16) of a_j.

Lemma 12.4. *For $n \to \infty$,*

$$\frac{\mathbf{E}\left(\widetilde{R}_n \otimes \widetilde{R}_n \mid \underline{\eta}\right)}{n \log n} = \frac{\sum_{j=1}^{n} a_j^2 U_{j-1} \otimes U_{j-1}}{n \log n}$$

$$\xrightarrow{\mathbf{P}} \sigma_d^2 \begin{pmatrix} t_1 I_d & & & \\ & (t_2 - t_1) I_d & & \\ & & \ddots & \\ & & & (t_k - t_{k-1}) I_d \end{pmatrix}. \tag{12.11}$$

This follows from Lemma 4.3 by observing that the different d-dimensional components of $\widetilde{R}_n$ are independent when conditioned on $\underline{\eta}$ and v_0. Note that the variance A_n defined in (4.20) satisfies $A_n^2 = \mathbf{E}(\|\widetilde{R}_n\|^2)$.

Given these lemmas, let us now conclude the proof of (12.5). The sequence of random vectors

$$Z_n := \frac{\widetilde{R}_n}{\sigma_d \sqrt{n \log n}} \tag{12.12}$$

is tight in $\mathbb{R}^{kd}$ because each component is tight in $\mathbb{R}^d$; the latter was proved at the end of Sect. 4. By the Helly–Prokhorov theorem, there is an infinite subset $S_1 \subset \mathbb{N}$ so that Z_n converges in distribution along $n \in S_1$ to some limit Z. Assume for a contradiction that Z is *not* the right hand side of (12.5). The Borel–Cantelli lemma implies that there is an infinite subset $S_2 \subset S_1$, so that in the statements of Lemmas 12.4 and 4.4 we have almost-sure convergence along $n \in S_2$:

$$\mathbf{E}\left(Z_n \otimes Z_n \mid \underline{\eta}\right) \xrightarrow{\text{a.s.}} \begin{pmatrix} t_1 I_d & & & \\ & (t_2 - t_1) I_d & & \\ & & \ddots & \\ & & & (t_k - t_{k-1}) I_d \end{pmatrix}, \tag{12.13}$$

$$\frac{A_n^2}{n \log n} \xrightarrow{\text{a.s.}} d \sigma_d^2, \tag{12.14}$$

and

$$A_n^{-2} \sum_{j=1}^{n} \mathbf{E}\big(\tilde{\xi}_j^2 \mathbb{1}_{\{\tilde{\xi}_j^2 > \varepsilon^2 A_n^2\}} \mid \boldsymbol{\eta}\big) \xrightarrow{\text{a.s.}} 0. \tag{12.15}$$

The hypotheses of the Lindeberg central limit theorem are thus satisfied, and we conclude the proof as for the one-dimensional distribution (see end of Sect. 4) with Lemmas 4.1 and 4.2 replaced by Lemmas 12.2 and 12.3. $\square$

13. Tightness

We will now establish tightness for the sequence of processes $(Y_n)_{n=1}^{\infty}$. This is the last remaining input in the proof of Theorem 1.4. We assume for simplicity that the scattering map satisfies the hypotheses of Sect. 2 (as opposed to the milder condition in Theorem 3.2). Define

$$\xi_{j,n} := \xi_j \mathbb{1}_{\{\xi_j \le r_n\}}, \qquad r_n = \sqrt{n (\log n)^\gamma}, \tag{13.1}$$

$$\tilde{\xi}_{j,n} = \xi_{j,n} - m_{j,n}, \tag{13.2}$$

with the conditional expectation

$$m_{j,n} := \mathbf{E}\left(\xi_{j,n} \mid \boldsymbol{\eta}\right) = \frac{K_{1,r_n}(\boldsymbol{\eta}_{j-1}, \boldsymbol{\eta}_j)}{K_0(\boldsymbol{\eta}_{j-1}, \boldsymbol{\eta}_j)}. \tag{13.3}$$

Furthermore, let

$$\boldsymbol{Q}_n^* := \sum_{j=1}^{n} \xi_{j,n} \boldsymbol{V}_{j-1}, \tag{13.4}$$

$$a_{j,n}^2 := \mathbf{Var}\left(\xi_{j,n} \mid \boldsymbol{\eta}\right) = \frac{K_{2,r_n}(\boldsymbol{\eta}_{j-1}, \boldsymbol{\eta}_j)}{K_0(\boldsymbol{\eta}_{j-1}, \boldsymbol{\eta}_j)} - m_{j,n}^2, \tag{13.5}$$

and

$$\mathcal{A}_n^2 := \sum_{j=1}^{n} a_{j,n}^2 = \mathbf{E}\left(\| \boldsymbol{Q}_n^* \|^2 \mid \boldsymbol{\eta}\right). \tag{13.6}$$

We split the process Y_n defined in (4.22) into four parts,

$$Y_n = \widehat{Y}_n + \widetilde{Y}_n + \overline{Y}_n + \check{Y}_n, \tag{13.7}$$

where

$$\widehat{Y}_n(t) = \frac{1}{\sigma_d \sqrt{n \log n}} \sum_{j=1}^{\lfloor nt \rfloor} \xi_j \mathbb{1}_{\{\xi_j > r_n\}} \boldsymbol{V}_{j-1}, \tag{13.8}$$

$$\widetilde{Y}_n(t) = \frac{1}{\sigma_d \sqrt{n \log n}} \sum_{j=1}^{\lfloor nt \rfloor} \tilde{\xi}_{j,n} \boldsymbol{V}_{j-1}, \tag{13.9}$$

$$\overline{Y}_n(t) = \frac{1}{\sigma_d \sqrt{n \log n}} \sum_{j=1}^{\lfloor nt \rfloor} m_{j,n} \boldsymbol{V}_{j-1}, \tag{13.10}$$

and

$$\check{\boldsymbol{Y}}_n(t) = \frac{1}{\sigma_d \sqrt{n \log n}} \{nt\} \, \xi_{\lfloor nt \rfloor + 1} \boldsymbol{V}_{\lfloor nt \rfloor}. \tag{13.11}$$

We begin by showing that $\widehat{\boldsymbol{Y}}_n$, $\overline{\boldsymbol{Y}}_n$ and $\check{\boldsymbol{Y}}_n$ are uniformly small with large probability. Consider first $\widehat{\boldsymbol{Y}}_n$ and $\check{\boldsymbol{Y}}_n$.

Proposition 13.1. *We have*

$$\sup_{n \in \mathbb{N}} \sup_{t \in [0,1]} \sqrt{n \log n} \, \|\widehat{\boldsymbol{Y}}_n(t)\| < \infty \tag{13.12}$$

almost surely.

The proof of Proposition 13.1 is identical to that of Lemma 4.1. The key ingredient is the Borel–Cantelli Lemma, cf. the comment after Lemma 12.2.

Proposition 13.2. *There is $C < \infty$ such that, for any $\beta > 0$ and any $n \geq 2$,*

$$\mathbf{P} \left(\sup_{t \in [0,1]} \|\check{\boldsymbol{Y}}_n(t)\| \geq \beta \right) \leq \frac{C}{\beta^2 \log n}. \tag{13.13}$$

Proof. We have

$$\mathbf{P} \left(\sup_{t \in [0,1]} \|\check{\boldsymbol{Y}}_n(t)\| \geq \beta \right) \leq \mathbf{P} \left(\max_{1 \leq j \leq n+1} |\xi_j| \geq \beta \sigma_d \sqrt{n \log n} \right)$$
$$\leq \sum_{j=1}^{n+1} \mathbf{P} \left(|\xi_j| \geq \beta \sigma_d \sqrt{n \log n} \right) \tag{13.14}$$

where, by the asymptotic tail for the free path length distribution [22, Theorem 1.14], we have for all $j \geq 1$

$$\mathbf{P} \left(|\xi_j| \geq \beta \sigma_d \sqrt{n \log n} \right) = O \left(\frac{1}{\beta^2 n \log n} \right). \tag{13.15}$$

$\square$

The estimation of $\overline{\boldsymbol{Y}}_n(t)$ relies on the following maximal inequality for martingales with stationary increments, cf. Gordin and Lifsic [18]. Let $(\mathcal{V}, \mu)$ be a probability space. We denote by $\mathrm{L}_0^2(\mathcal{V}, \mu)$ the orthogonal complement of the constant functions in $\mathrm{L}^2(\mathcal{V}, \mu)$.

Proposition 13.3. *Let $\underline{\boldsymbol{\alpha}} = (\boldsymbol{\alpha}_n)_{n=0}^\infty$ be a Markov chain on the state space $\mathcal{V}$, and μ a probability measure which is stationary and ergodic for $\underline{\boldsymbol{\alpha}}$. Let $\mathcal{P}$ be the transition operator on $\mathrm{L}^2(\mathcal{V}, \mu)$ defined by $\mathcal{P}f(z) = \mathbf{E}\left(f(\boldsymbol{\alpha}_n) \mid \boldsymbol{\alpha}_{n-1} = z \right)$. Then, for any $f \in \mathrm{Ran}(\mathcal{I} - \mathcal{P})$, $n \in \mathbb{N}$, $\kappa > 0$,*

$$\mathbf{P} \left(\max_{1 \leq m \leq n} \left| \sum_{j=1}^m f(\boldsymbol{\alpha}_j) \right| \geq \kappa \sqrt{n} \right) \leq \frac{9}{\kappa^2} \left(\|g\|^2 + \frac{1}{n} \|\mathcal{P}g\|^2 \right), \tag{13.16}$$

where g is the unique function in $\mathrm{L}_0^2(\mathcal{V}, \mu)$ such that $f = (\mathcal{I} - \mathcal{P})g$.

Proof. We have

$$\sum_{j=1}^{m} f(\boldsymbol{\alpha}_j) = \sum_{j=1}^{m} \big(g(\boldsymbol{\alpha}_j) - \mathcal{P}g(\boldsymbol{\alpha}_j)\big) = \mathcal{M}_m + \mathcal{P}g(\boldsymbol{\alpha}_0) - \mathcal{P}g(\boldsymbol{\alpha}_m). \quad (13.17)$$

where

$$\mathcal{M}_m := \sum_{j=1}^{m} \big(g(\boldsymbol{\alpha}_j) - \mathcal{P}g(\boldsymbol{\alpha}_{j-1})\big) \quad (13.18)$$

is a martingale with stationary and ergodic increments. The left hand side of (13.16) is estimated by the sum of the following three terms. The first is bounded by Doob's inequality for non-negative sub-martingales,

$$\mathbf{P}\left(\max_{1\leq m\leq n} |\mathcal{M}_m| \geq \frac{\kappa\sqrt{n}}{3}\right) \leq \frac{9}{\kappa^2 n} \mathbf{E}\left(|\mathcal{M}_n|^2\right) = \frac{9}{\kappa^2}\big(\|g\|^2 - \|\mathcal{P}g\|^2\big). \quad (13.19)$$

The second follows from Chebyshev's inequality,

$$\mathbf{P}\left(|\mathcal{P}g(\boldsymbol{\alpha}_0)| \geq \frac{\kappa\sqrt{n}}{3}\right) \leq \frac{9}{\kappa^2 n}\|\mathcal{P}g\|^2, \quad (13.20)$$

and the third from the union bound and Chebyshev's inequality,

$$\mathbf{P}\left(\max_{1\leq m\leq n} |\mathcal{P}g(\boldsymbol{\alpha}_m)| \geq \frac{\kappa\sqrt{n}}{3}\right) \leq \sum_{m=1}^{n} \mathbf{P}\left(|\mathcal{P}g(\boldsymbol{\alpha}_m)| \geq \frac{\kappa\sqrt{n}}{3}\right) \leq \frac{9}{\kappa^2}\|\mathcal{P}g\|^2.$$
$$(13.21)$$

$\square$

Proposition 13.4. *There is $C < \infty$ such that, for any $\beta > 0$ and any $n \geq 2$,*

$$\mathbf{P}\left(\sup_{t\in[0,1]} \|\overline{\boldsymbol{Y}}_n(t)\| \geq \beta\right) \leq \begin{cases} \dfrac{C \log\log n}{\beta^2 \log n} & (d=2) \\ \dfrac{C}{\beta^2 \log n} & (d \geq 3). \end{cases} \quad (13.22)$$

Proof. The plan is to apply Proposition 13.3 to a Markov chain defined on the state space of three consecutive velocities,

$$\mathcal{V} := \{(\boldsymbol{v}_{n-1}, \boldsymbol{v}_n, \boldsymbol{v}_{n+1}) \in (\mathrm{S}_1^{d-1})^3 : \varphi(\boldsymbol{v}_{n-1}, \boldsymbol{v}_n) > B_\theta, \ \varphi(\boldsymbol{v}_n, \boldsymbol{v}_{n+1}) > B_\theta\}, \quad (13.23)$$

where $\varphi(\boldsymbol{u}_1, \boldsymbol{u}_2) \in [0, \pi]$ denotes the angle between the two vectors $\boldsymbol{u}_1, \boldsymbol{u}_2$, and B_θ as in (2.4). For $(\boldsymbol{v}_{n-1}, \boldsymbol{v}_n, \boldsymbol{v}_{n+1}) \in \mathcal{V}$ and $\xi > 0$, let

$$p_0(\boldsymbol{v}_{n-1}, \boldsymbol{v}_n, \xi, \boldsymbol{v}_{n+1}) = \Psi_0\big(\boldsymbol{w}(\boldsymbol{v}_{n-1}, \boldsymbol{v}_n), \xi, \boldsymbol{w}'(\boldsymbol{v}_n, \boldsymbol{v}_{n+1})\big) \sigma(\boldsymbol{v}_n, \boldsymbol{v}_{n+1}) \quad (13.24)$$

where the functions $\boldsymbol{w}, \boldsymbol{w}'$ are defined via

$$\begin{pmatrix} 0 \\ \boldsymbol{w}(\boldsymbol{v}_{n-1}, \boldsymbol{v}_n) \end{pmatrix} = R(\boldsymbol{v}_n)^{-1}\boldsymbol{s}(\boldsymbol{v}_{n-1}, \boldsymbol{v}_n), \quad (13.25)$$

$$\begin{pmatrix} 0 \\ \boldsymbol{w}'(\boldsymbol{v}_n, \boldsymbol{v}_{n+1}) \end{pmatrix} = R(\boldsymbol{v}_n)^{-1}\boldsymbol{b}(\boldsymbol{v}_n, \boldsymbol{v}_{n+1}); \quad (13.26)$$

here $s(\boldsymbol{v}_{n-1}, \boldsymbol{v}_n)$, $\boldsymbol{b}(\boldsymbol{v}_n, \boldsymbol{v}_{n+1}) \in \mathbb{R}^d$ are the exit and impact parameters, respectively, expressed as functions of incoming and outgoing velocities. Furthermore $\sigma(\boldsymbol{v}_0, \boldsymbol{v})$ is the differential cross section, i.e. the differential of the map $\boldsymbol{v} \mapsto \boldsymbol{b}(\boldsymbol{v}_0, \boldsymbol{v})$. The function $p_0(\boldsymbol{v}_{n-1}, \boldsymbol{v}_n, \xi, \boldsymbol{v}_{n+1})$ is precisely the transition kernel that governs the Boltzmann–Grad limit of the periodic Lorentz gas in the velocity representation used in [21]. We integrate out the flight time and obtain

$$\int_0^\infty p_0(\boldsymbol{v}_{n-1}, \boldsymbol{v}_n, \xi, \boldsymbol{v}_{n+1}) \, d\xi = L_0(\boldsymbol{v}_{n-1}, \boldsymbol{v}_n, \boldsymbol{v}_{n+1}) \, \sigma(\boldsymbol{v}_n, \boldsymbol{v}_{n+1}) \quad (13.27)$$

with

$$L_0(\boldsymbol{v}_{n-1}, \boldsymbol{v}_n, \boldsymbol{v}_{n+1}) := K_0\big(\boldsymbol{w}(\boldsymbol{v}_{n-1}, \boldsymbol{v}_n), \boldsymbol{w}'(\boldsymbol{v}_n, \boldsymbol{v}_{n+1})\big), \quad (13.28)$$

with the kernel $K_0(\boldsymbol{w}, z)$ defined in (4.5). Thus

$$n \mapsto \boldsymbol{\alpha}_n = (V_{n-1}, V_n, V_{n+1}) \quad (13.29)$$

[where V_j is the random variable defined in (3.10)] defines a Markov chain on the state space $\mathcal{V}$ with stationary measure

$$d\mu((\boldsymbol{v}_{n-1}, \boldsymbol{v}_n, \boldsymbol{v}_{n+1})) := \bar{s}^{-1}\bar{\sigma}^{-2} \, \sigma(\boldsymbol{v}_{n-1}, \boldsymbol{v}_n) \, \sigma(\boldsymbol{v}_n, \boldsymbol{v}_{n+1}) \, d\boldsymbol{v}_{n-1} \, d\boldsymbol{v}_n \, d\boldsymbol{v}_{n+1} \quad (13.30)$$

where $\bar{s}$ is the volume of S_1^{d-1} and $\bar{\sigma}$ the volume of $\mathcal{B}_1^{d-1}$. Explicitly, for $\mathcal{A} \subset \mathcal{V}$,

$$\mathbf{P}\big(\boldsymbol{\alpha}_n \in \mathcal{A} \mid \boldsymbol{\alpha}_{n-1} = z\big) = \int_{\mathcal{A}} \mathcal{K}(z, z') \, d\mu(z') \quad (13.31)$$

with kernel

$$\mathcal{K}\big((\boldsymbol{v}_{n-2}, \boldsymbol{v}_{n-1}, \boldsymbol{v}_n), (\boldsymbol{v}'_{n-1}, \boldsymbol{v}'_n, \boldsymbol{v}'_{n+1})\big)$$
$$= \bar{s}\,\bar{\sigma}^2 \, \frac{\delta(\boldsymbol{v}_{n-1}, \boldsymbol{v}'_{n-1}) \, \delta(\boldsymbol{v}_n, \boldsymbol{v}'_n) \, L_0(\boldsymbol{v}'_{n-1}, \boldsymbol{v}'_n, \boldsymbol{v}'_{n+1})}{\sigma(\boldsymbol{v}_{n-1}, \boldsymbol{v}_n)}. \quad (13.32)$$

Note that the kernel $\mathcal{K}(z, z')$ is defined with respect to the measure μ and *not* with respect to the Lebesgue measure.

The transition operator $\mathcal{P}$ for this Markov chain is defined by

$$\mathcal{P}f(z) = \mathbf{E}\big(f(\boldsymbol{\alpha}_n) \mid \boldsymbol{\alpha}_{n-1} = z\big)$$
$$= \int \mathcal{K}(z, z') f(z') \, d\mu(z'). \quad (13.33)$$

The operator $\mathcal{P}$ has eigenvalue 1 (corresponding to constant eigenfunctions). To establish that there is a spectral gap, we calculate the kernel $\mathcal{K}^{(m)}(z, z')$ of the mth power,

$$\mathcal{P}^m f(z) = \mathbf{E}\big(f(\boldsymbol{\alpha}_n) \mid \boldsymbol{\alpha}_{n-1} = z\big)$$
$$= \int \mathcal{K}^{(m)}(z, z') f(z') \, d\mu(z'). \quad (13.34)$$

The second power reads

$$\mathcal{K}^{(2)}\big((\boldsymbol{v}_{n-2}, \boldsymbol{v}_{n-1}, \boldsymbol{v}_n), (\boldsymbol{v}'_n, \boldsymbol{v}'_{n+1}, \boldsymbol{v}'_{n+2})\big)$$
$$= \bar{s}\,\bar{\sigma}^2 L_0(\boldsymbol{v}_{n-1}, \boldsymbol{v}_n, \boldsymbol{v}'_{n+1}) \, L_0(\boldsymbol{v}'_n, \boldsymbol{v}'_{n+1}, \boldsymbol{v}'_{n+2}) \, \delta(\boldsymbol{v}_n, \boldsymbol{v}'_n). \quad (13.35)$$

As to the third power,

$$
\begin{aligned}
&\mathcal{K}^{(3)}\big((\boldsymbol{v}_{n-2}, \boldsymbol{v}_{n-1}, \boldsymbol{v}_n), (\boldsymbol{v}'_{n+1}, \boldsymbol{v}'_{n+2}, \boldsymbol{v}'_{n+3})\big) \\
&= \bar{s}\,\bar{\sigma}^2 L_0(\boldsymbol{v}_{n-1}, \boldsymbol{v}_n, \boldsymbol{v}'_{n+1})\, L_0(\boldsymbol{v}_n, \boldsymbol{v}'_{n+1}, \boldsymbol{v}'_{n+2})\, L_0(\boldsymbol{v}'_{n+1}, \boldsymbol{v}'_{n+2}, \boldsymbol{v}'_{n+3})\, \sigma(\boldsymbol{v}_n, \boldsymbol{v}'_{n+1}).
\end{aligned}
\tag{13.36}
$$

In view of the lower bound on $K_0(\boldsymbol{w}, \boldsymbol{z})$ (Lemmas 6.1, 6.2), we have

$$
\mathcal{K}^{(3)}\big((\boldsymbol{v}_{n-2}, \boldsymbol{v}_{n-1}, \boldsymbol{v}_n), (\boldsymbol{v}'_{n+1}, \boldsymbol{v}'_{n+2}, \boldsymbol{v}'_{n+3})\big) \geq \frac{\bar{s}\,\bar{\sigma}^2}{(2^d \bar{\sigma} \zeta(d))^3}\, \sigma(\boldsymbol{v}_n, \boldsymbol{v}'_{n+1}). \tag{13.37}
$$

The fourth power reads

$$
\begin{aligned}
&\mathcal{K}^{(4)}\big((\boldsymbol{v}_{n-2}, \boldsymbol{v}_{n-1}, \boldsymbol{v}_n), (\boldsymbol{v}'_{n+2}, \boldsymbol{v}'_{n+3}, \boldsymbol{v}'_{n+4})\big) \\
&= \bar{s}\,\bar{\sigma}^2 \int_{S_1^{d-1}} L_0(\boldsymbol{v}_{n-1}, \boldsymbol{v}_n, \tilde{\boldsymbol{v}}_{n+1})\, L_0(\boldsymbol{v}_n, \tilde{\boldsymbol{v}}_{n+1}, \boldsymbol{v}'_{n+2})\, L_0(\tilde{\boldsymbol{v}}_{n+1}, \boldsymbol{v}'_{n+2}, \boldsymbol{v}'_{n+3}) \\
&\quad \times L_0(\boldsymbol{v}'_{n+2}, \boldsymbol{v}'_{n+3}, \boldsymbol{v}'_{n+4})\, \sigma(\boldsymbol{v}_n, \tilde{\boldsymbol{v}}_{n+1})\, \sigma(\tilde{\boldsymbol{v}}_{n+1}, \boldsymbol{v}'_{n+2})\, d\tilde{\boldsymbol{v}}_{n+1}.
\end{aligned}
\tag{13.38}
$$

The lower bound on $K_0(\boldsymbol{w}, \boldsymbol{z})$ now yields

$$
\begin{aligned}
&\mathcal{K}^{(4)}\big((\boldsymbol{v}_{n-2}, \boldsymbol{v}_{n-1}, \boldsymbol{v}_n), (\boldsymbol{v}'_{n+2}, \boldsymbol{v}'_{n+3}, \boldsymbol{v}'_{n+4})\big) \\
&\geq \frac{\bar{s}\,\bar{\sigma}^2}{(2^d \bar{\sigma} \zeta(d))^4} \int_{S_1^{d-1}} \sigma(\boldsymbol{v}_n, \tilde{\boldsymbol{v}}_{n+1})\, \sigma(\tilde{\boldsymbol{v}}_{n+1}, \boldsymbol{v}'_{n+2})\, d\tilde{\boldsymbol{v}}_{n+1}.
\end{aligned}
\tag{13.39}
$$

Similarly, for the mth power ($m \geq 5$),

$$
\begin{aligned}
&\mathcal{K}^{(m)}\big((\boldsymbol{v}_{n-2}, \boldsymbol{v}_{n-1}, \boldsymbol{v}_n), (\boldsymbol{v}'_{n+m-2}, \boldsymbol{v}'_{n+m-1}, \boldsymbol{v}'_{n+m})\big) \\
&\geq \frac{\bar{s}\,\bar{\sigma}^2}{(2^d \bar{\sigma} \zeta(d))^m} \int_{(S_1^{d-1})^{m-3}} \sigma(\boldsymbol{v}_n, \tilde{\boldsymbol{v}}_{n+1})\, \sigma(\tilde{\boldsymbol{v}}_{n+1}, \tilde{\boldsymbol{v}}_{n+2}) \cdots \\
&\qquad \cdots \sigma(\tilde{\boldsymbol{v}}_{n+m-4}, \tilde{\boldsymbol{v}}_{n+m-3})\, \sigma(\tilde{\boldsymbol{v}}_{n+m-3}, \boldsymbol{v}'_{n+m-2})\, d\tilde{\boldsymbol{v}}_{n+1} \cdots d\tilde{\boldsymbol{v}}_{n+m-3}. \tag{13.40}
\end{aligned}
$$

For the class of scattering maps considered in Sect. 2, there exists a finite m such that the integral on the right-hand side has a uniform positive lower bound for all $\boldsymbol{v}_n, \boldsymbol{v}'_{n+m-2} \in S_1^{d-1}$. Hence

$$
\inf_{\boldsymbol{z}, \boldsymbol{z}' \in \mathcal{V}} \mathcal{K}^{(m)}(\boldsymbol{z}, \boldsymbol{z}') > 0. \tag{13.41}
$$

This, by the classic Doeblin argument (recall Sect. 7), implies that $\mathcal{P}^m$, and therefore $\mathcal{P}$, has a spectral gap.

Let

$$
L_{1,r}(\boldsymbol{v}_{n-1}, \boldsymbol{v}_n, \boldsymbol{v}_{n+1}) := K_{1,r}\big(\boldsymbol{w}(\boldsymbol{v}_{n-1}, \boldsymbol{v}_n), \boldsymbol{w}'(\boldsymbol{v}_n, \boldsymbol{v}_{n+1})\big), \tag{13.42}
$$

with $K_{1,r}(\boldsymbol{w}, \boldsymbol{z})$ as in (4.12), and define the function $\ell_r : \mathcal{V} \to \mathbb{R}_{\geq 0}$ by

$$
\ell_r((\boldsymbol{v}_{n-1}, \boldsymbol{v}_n, \boldsymbol{v}_{n+1})) := \frac{L_{1,r}(\boldsymbol{v}_{n-1}, \boldsymbol{v}_n, \boldsymbol{v}_{n+1})}{L_0(\boldsymbol{v}_{n-1}, \boldsymbol{v}_n, \boldsymbol{v}_{n+1})}. \tag{13.43}
$$

We then recover the random variable (13.3) via

$$m_{j,n} = \ell_{r_n}(\boldsymbol{\alpha}_n) = \ell_{r_n}((\boldsymbol{V}_{n-1}, \boldsymbol{V}_n, \boldsymbol{V}_{n+1})). \tag{13.44}$$

To conclude the proof, apply Proposition 13.3 with the (n-dependent) choice of $f \in L^2(\mathcal{V}, d\mu)$,

$$f((\boldsymbol{v}_{n-1}, \boldsymbol{v}_n, \boldsymbol{v}_{n+1})) = \boldsymbol{e} \cdot \boldsymbol{v}_n \, \ell_{r_n}((\boldsymbol{v}_{n-1}, \boldsymbol{v}_n, \boldsymbol{v}_{n+1})) \tag{13.45}$$

with an arbitrary fixed $\boldsymbol{e} \in S_1^{d-1}$. Since $\ell_{r_n}((R\boldsymbol{v}_{n-1}, R\boldsymbol{v}_n, R\boldsymbol{v}_{n+1})) = \ell_{r_n}((\boldsymbol{v}_{n-1}, \boldsymbol{v}_n, \boldsymbol{v}_{n+1}))$ and $d\mu((R\boldsymbol{v}_{n-1}, R\boldsymbol{v}_n, R\boldsymbol{v}_{n+1})) = d\mu((\boldsymbol{v}_{n-1}, \boldsymbol{v}_n, \boldsymbol{v}_{n+1}))$ for all $R \in SO(d)$, we have

$$\int_{\mathcal{V}} f(z) \, d\mu(z) = 0, \tag{13.46}$$

i.e., $f \in L_0^2(\mathcal{V}, \mu)$. Thanks to the spectral gap of $\mathcal{P}$, there is a constant $M < \infty$, such that for all $f \in L_0^2(\mathcal{V}, \mu)$,

$$\|(\mathcal{I} - \mathcal{P})^{-1} f\|^2 \le M \|f\|^2 \tag{13.47}$$

Finally, the estimate for $\mathbf{E}(m_n^2)$ in the proof of Lemma 4.2 yields $\|f\|^2 = O(\log\log n)$ if $d = 2$ and $\|f\|^2 = O(1)$ if $d \ge 3$. Proposition 13.4 thus follows from Proposition 13.3 with $\kappa = \beta\sqrt{\log n}$. $\quad\square$

Propositions 13.1, 13.2 and 13.4 establish that the tightness for $(Y_n)_{n=1}^{\infty}$ is implied by the tightness for $(\widetilde{Y}_n)_{n=1}^{\infty}$. To prove the latter, we use the following classical characterization of tightness for a random curve X.

Theorem 13.5 [5, Theorem 8.3]. *The sequence* $(\lambda_n)_{n=1}^{\infty}$ *of probability measures in* $C_0([0, 1])$ *is tight if for every* $\epsilon > 0$, $\beta > 0$ *there exist* $\delta < 1$, $n_0 < \infty$ *such that for all* $n \ge n_0$ *we have*

$$\frac{1}{\delta} \sup_{t \in [0,1]} \lambda_n \left(\sup_{s \in [t, t+\delta] \cap [0,1]} \|X(s) - X(t)\| \ge \beta \right) < \epsilon. \tag{13.48}$$

We will also exploit the following maximal inequality for sums of independent random variables.

Lemma 13.6 [5, Lemma, p. 69]. *Let* $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_m$ *be independent random variables in* $\mathbb{R}^d$ *with mean zero and finite variances* $\sigma_i^2 = \mathbf{E}(\|\boldsymbol{\xi}_i\|^2)$. *Put* $S_m = \boldsymbol{\xi}_1 + \ldots + \boldsymbol{\xi}_m$ *and* $s_m^2 = \sigma_1^2 + \ldots + \sigma_m^2$. *Then, for any* $\lambda \in \mathbb{R}$,

$$\mathbf{P} \left(\max_{i \le m} \|S_i\| \ge \lambda s_m \right) \le 2 \mathbf{P} \left(\|S_m\| \ge (\lambda - \sqrt{2}) s_m \right). \tag{13.49}$$

The following proposition verifies the hypothesis of Theorem 13.5, and thus proves that the sequence of probability measures corresponding to $(\widetilde{Y}_n)_{n=1}^{\infty}$ is tight.

Proposition 13.7. *For every* $\epsilon > 0$, $\beta > 0$ *there exist* $\delta < 1$, $n_0 < \infty$ *such that for all* $n \ge n_0$ *we have*

$$\frac{1}{\delta} \sup_{t \in [0,1]} \mathbf{P} \left(\sup_{s \in [t, t+\delta] \cap [0,1]} \|\widetilde{Y}_n(s) - \widetilde{Y}_n(t)\| \ge \beta \right) < \epsilon. \tag{13.50}$$

Proof. We need to show that, for every $\epsilon > 0$, $\beta > 0$ there exist $\delta < 1$, $n_0 < \infty$ such that for all $n \geq n_0$ we have

$$\frac{1}{\delta} \mathbf{P} \left(\max_{nt < m \leq n(t+\delta)} \left\| \sum_{j=\lfloor nt \rfloor + 1}^{m} \tilde{\xi}_{j,n} V_{j-1} \right\| \geq \beta \sqrt{n \log n} \right) < \epsilon. \tag{13.51}$$

To prove this fact, note first that

$$\mathbf{P} \left(\max_{nt < m \leq n(t+\delta)} \left\| \sum_{j=\lfloor nt \rfloor + 1}^{m} \tilde{\xi}_{j,n} V_{j-1} \right\| \geq \beta \sqrt{n \log n} \right)$$

$$= \mathbf{E} \left(\mathbf{P} \left(\max_{nt < m \leq n(t+\delta)} \left\| \sum_{j=\lfloor nt \rfloor + 1}^{m} \tilde{\xi}_{j,n} V_{j-1} \right\| \geq \beta \sqrt{n \log n} \; \Big| \; \underline{\eta} \right) \right). \tag{13.52}$$

The maximal inequality (13.49) yields

$$\mathbf{P} \left(\max_{nt < m \leq n(t+\delta)} \left\| \sum_{j=\lfloor nt \rfloor + 1}^{m} \tilde{\xi}_{j,n} V_{j-1} \right\| \geq \beta \sqrt{n \log n} \; \Big| \; \underline{\eta} \right)$$

$$\leq 2 \mathbf{P} \left(\left\| \sum_{j=\lfloor nt \rfloor + 1}^{\lfloor n(t+\delta) \rfloor} \tilde{\xi}_{j,n} V_{j-1} \right\| \geq \beta \sqrt{n \log n} - \sqrt{2(\mathcal{A}_{\lfloor n(t+\delta) \rfloor}^2 - \mathcal{A}_{\lfloor nt \rfloor}^2)} \; \Big| \; \underline{\eta} \right), \tag{13.53}$$

which in turn implies, after taking expectation values,

$$\mathbf{P} \left(\max_{nt < m \leq n(t+\delta)} \left\| \sum_{j=\lfloor nt \rfloor + 1}^{m} \tilde{\xi}_{j,n} V_{j-1} \right\| \geq \beta \sqrt{n \log n} \right)$$

$$\leq 2 \mathbf{P} \left(\left\| \sum_{j=\lfloor nt \rfloor + 1}^{\lfloor n(t+\delta) \rfloor} \tilde{\xi}_{j,n} V_{j-1} \right\| \geq \beta \sqrt{n \log n} - \sqrt{2(\mathcal{A}_{\lfloor n(t+\delta) \rfloor}^2 - \mathcal{A}_{\lfloor nt \rfloor}^2)} \right). \tag{13.54}$$

The latter is bounded above by

$$\leq 2 \mathbf{P} \left(\left\| \sum_{j=\lfloor nt \rfloor + 1}^{\lfloor n(t+\delta) \rfloor} \tilde{\xi}_{j,n} V_{j-1} \right\| \geq (\beta - K\sqrt{\delta}) \sqrt{n \log n} \right)$$

$$+ 2 \mathbf{P} \left(\mathcal{A}_{\lfloor n(t+\delta) \rfloor}^2 - \mathcal{A}_{\lfloor nt \rfloor}^2 > K^2 \delta n \log n \right), \tag{13.55}$$

for any constant $K \geq 0$. By adapting the proof of Lemma 4.3 (replacing $j(\log j)^\gamma$ by $n(\log n)^\gamma$), one can prove that the analogue of (4.19) reads

$$\frac{\mathcal{A}_n^2}{n \log n} \xrightarrow{\mathbf{P}} d \sigma_d^2. \tag{13.56}$$

This implies that for any constant $K > \sigma_d \sqrt{d}$ we have

$$\lim_{n \to \infty} \mathbf{P} \left(\mathcal{A}_{\lfloor n(t+\delta) \rfloor}^2 - \mathcal{A}_{\lfloor nt \rfloor}^2 > K^2 \delta n \log n \right) = 0 \tag{13.57}$$

uniformly in t, δ, which takes care of the second term in (13.55). The first term in (13.55) is estimated by Theorem 3.2 (ii): Given β, ϵ, for any sufficiently small δ there is n_0 such that for all $t \in [0, 1]$ and $n \geq n_0$

$$
\mathbf{P}\left(\left\|\sum_{j=\lfloor nt \rfloor+1}^{\lfloor n(t+\delta) \rfloor} \tilde{\xi}_{j,n} \mathbf{V}_{j-1}\right\| \geq \left(\beta - K\sqrt{\delta}\right)\sqrt{n \log n}\right)
$$

$$
\leq \frac{1}{(2\pi)^{d/2}} \int_{\|\mathbf{x}\| > \frac{\beta - K\sqrt{\delta}}{\sigma_d \sqrt{\delta}}} e^{-\frac{1}{2}\|\mathbf{x}\|^2} d\mathbf{x} + \epsilon \delta = O(\epsilon \delta). \tag{13.58}
$$

Note that here we have applied Theorem 3.2 (ii) to the truncated $\tilde{\xi}_{j,n}$ rather than ξ_j, which is justified by the analogue of Lemmas 4.1, 4.2 with $j(\log j)^\gamma$ replaced by $n(\log n)^\gamma$. This completes the proof. $\quad\square$

14. Theorem 1.4 Implies Theorem 1.3

We now turn to the continuous time process. The Boltzmann–Grad limit $r \to 0$ is covered by [21, Theorem 1.2], which tells us that for arbitrary fixed T,

$$
X_{T,r} \Rightarrow X_T \tag{14.1}
$$

where

$$
X_T(t) = \frac{X(tT)}{\Sigma_d \sqrt{T \log T}} = \frac{\mathbf{Q}_{v_{tT}} + (t - \tau_{v_{tT}})\mathbf{V}_{v_{tT}}}{\Sigma_d \sqrt{T \log T}}; \tag{14.2}
$$

recall (3.8)–(3.12).

The convergence of finite-dimensional distributions of X_T to W follows (within the framework of Sect. 12) from the same estimates as in Sect. 11. What remains is to show tightness in C([0, 1]) for the family of processes $(X_T)_{T \geq 1}$. By a simple scaling argument, it is sufficient to prove tightness for the sequence $(X_n)_{n \in \mathbb{N}}$.

Define the continuous, strictly increasing (random) functions T, $\Theta : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ by

$$
T(\theta) := \tau_{\lfloor \theta \rfloor} + \{\theta\}\xi_{\lfloor \theta \rfloor+1}, \qquad \Theta(t) := v_t + \frac{t - \tau_{v_t}}{\xi_{v_t+1}}. \tag{14.3}
$$

Note that the functions T and Θ are the inverse of one another. That is, $T(\Theta(t)) = t$ and $\Theta(T(\theta)) = \theta$. The key point is that we can now relate the curves $t \to X_n(t)$ and $\theta \to \mathbf{Q}_n(\theta)$ by this time change: We have $X(t) = \mathbf{Q}_1(\Theta(t))$ and therefore

$$
X(nt) = \mathbf{Q}_1(\Theta(nt)) = \mathbf{Q}_n(n^{-1}\Theta(nt)) = \mathbf{Q}_n(\Theta_n(t)), \tag{14.4}
$$

where $\Theta_n(t) := n^{-1}\Theta(nt)$. This yields for the normalized processes

$$
X_n(t) = \bar{\xi}^{1/2} \mathbf{Y}_n(\Theta_n(t)), \tag{14.5}
$$

where $\bar{\xi} = 1/\bar{\sigma}$ is the mean free path length.

Given $b > 0$, consider the random process $Z_n : [0, b] \to \mathbb{R}$ defined by

$$
Z_n(\theta) := \frac{T(n\theta) - n\theta\bar{\xi}}{\sigma_d \sqrt{d n \log n}}. \tag{14.6}
$$

The following lemma says that Z_n converges to one-dimensional Brownian motion W.

Lemma 14.1. *For $n \to \infty$,*

$$Z_n \Rightarrow W. \tag{14.7}$$

The proof of this lemma is a simpler variant of the already established weak convergence $Y_n \Rightarrow W$.

The lemma implies in particular that the process $\theta \mapsto T_n(\theta) := n^{-1} T(n\theta)$ converges weakly to the deterministic function $\theta \mapsto \bar{\xi}\theta$. Since $T_n(\Theta_n(t)) = t$ and $\Theta_n(T_n(\theta)) = \theta$, this implies that $t \mapsto \Theta_n(t)$ converges weakly to $t \mapsto \bar{\sigma}t$.

The modulus of continuity of a curve $X \in C_0([0, b])$ is defined as

$$\omega_X^{[0,b]}(\delta) := \sup_{\substack{0 \le s, t \le b \\ |t-s| \le \delta}} \|X(s) - X(t)\|. \tag{14.8}$$

The tightness for $(X_n)_{n \in \mathbb{N}}$ is implied by the following lemma.

Lemma 14.2. *For every $\beta > 0$, $\epsilon > 0$ there exist $\delta < 1$ and $n_0 < \infty$ such that for all $n \ge n_0$*

$$\mathbf{P}\left(\omega_{X_n}^{[0,1]}(\delta) > \beta\right) < \epsilon. \tag{14.9}$$

Proof. Notice that

$$\mathbf{P}\left(\omega_{X_n}^{[0,1]}(\delta) > \beta\right) \le \mathbf{P}\left(\omega_{\bar{\xi}^{1/2} Y_n}^{[0,2\bar{\sigma}]} \circ \omega_{\Theta_n}^{[0,1]}(\delta) > \beta\right) + \mathbf{P}\left(\Theta_n(1+\delta) > 2\bar{\sigma}\right). \tag{14.10}$$

For $n \to \infty$, we have $\Theta_n(1+\delta) \xrightarrow{\mathbf{P}} (1+\delta)\bar{\sigma} < 2\bar{\sigma}$ (see the remark after Lemma 14.1), and therefore

$$\lim_{n \to \infty} \mathbf{P}\left(\Theta_n(1+\delta) > 2\bar{\sigma}\right) = 0. \tag{14.11}$$

The claim now follows from the tightness of $(Y_n)_{n=1}^{\infty}$ established in Theorem 1.4, and from the tightness of $(\Theta_n)_{n=1}^{\infty}$ which follows from the remark after Lemma 14.1. $\square$

This concludes the proof of Theorem 1.3.

References

1. Bálint, P., Chernov, N., Dolgopyat, D.: Limit theorems for dispersing billiards with cusps. Commun. Math. Phys. **308**(2), 479–510 (2011)
2. Bálint, P., Gouëzel, S.: Limit theorems in the stadium billiard. Commun. Math. Phys. **263**(2), 461–512 (2006)
3. Bálint, P., Tóth, I.P.: Exponential decay of correlations in multi-dimensional dispersing billiards. Annales Henri Poincaré J. Theor. Math. Phys. **9**(7), 1309–1369 (2008)
4. Bálint, P., Tóth, I.P.: Example for exponential growth of complexity in a finite horizon multi-dimensional dispersing billiard. Nonlinearity **25**(5), 1275–1297 (2012)

5. Billingsley, P.: Convergence of Probability Measures, 1st edition. Wiley, New York (1968). (Note that the second edition is not compatible with the references made)
6. Bleher, P.M.: Statistical properties of two-dimensional periodic Lorentz gas with infinite horizon. J. Stat. Phys. **66**(1–2), 315–373 (1992)
7. Boldrighini, C., Bunimovich, L.A., Sinai, Y.G.: On the Boltzmann equation for the Lorentz gas. J. Stat. Phys. **32**(3), 477–501 (1983)
8. Bourgain, J., Golse, F., Wennberg, B.: On the distribution of free path lengths for the periodic Lorentz gas. Commun. Math. Phys. **190**(3), 491–508 (1998)
9. Bunimovich, L.A., Sinai, Y.G.: Statistical properties of Lorentz gas with periodic configuration of scatterers. Commun. Math. Phys. **78**(4), 479–497 (1980)
10. Bykovskii, V.A., Ustinov, A.V.: The statistics of particle trajectories in the nonhomogeneous Sinai problem for a two-dimensional lattice. Izvestiya Math. **73**(4), 669–688 (2009)
11. Caglioti, E., Golse, F.: The Boltzmann–Grad limit of the periodic Lorentz gas in two space dimensions. C. R. Math. Acad. Sci. Paris **346**(7–8), 477–482 (2008)
12. Caglioti, E., Golse, F.: On the Boltzmann–Grad limit for the two dimensional periodic Lorentz gas. J. Stat. Phys. **141**(2), 264–317 (2010)
13. Chernov, N.I.: Statistical properties of the periodic Lorentz gas: multidimensional case. J. Stat. Phys. **74**(1–2), 11–53 (1994)
14. Dettmann, C.P.: New horizons in multidimensional diffusion: the Lorentz gas and the Riemann hypothesis. J. Stat. Phys. **146**(1), 181–204 (2012)
15. Dolgopyat, D.I., Chernov, N.I.: Anomalous current in periodic Lorentz gases with an infinite horizon. Russ. Math. Surv. **64**(4), 651–699 (2009)
16. Gallavotti, G.: Divergences and approach to equilibrium in the Lorentz and the wind-tree-models. Phys. Rev. **185**, 308–322 (1969)
17. Golse, F., Wennberg, B.: On the distribution of free path lengths for the periodic Lorentz gas. II. M2AN. Math. Model. Numer. Anal. **34**(6), 1151–1163 (2000)
18. Gordin, M.I., Lifsic, B.A.: Central limit theorem for stationary Markov processes. Sov. Math. Doklady **19**(2), 392–394 (1978)
19. Marklof, J., Strömbergsson, A.: Kinetic transport in the two-dimensional periodic Lorentz gas. Nonlinearity **21**(7), 1413–1422 (2008)
20. Marklof, J., Strömbergsson, A.: The distribution of free path lengths in the periodic Lorentz gas and related lattice point problems. Ann. Math. Second Ser. **172**(3), 1949–2033 (2010)
21. Marklof, J., Strömbergsson, A.: The Boltzmann–Grad limit of the periodic Lorentz gas. Ann. Math. Second Ser. **174**(1), 225–298 (2011)
22. Marklof, J., Strömbergsson, A.: The periodic Lorentz gas in the Boltzmann–Grad limit: asymptotic estimates. Geom. Funct. Anal. **21**(3), 560–647 (2011)
23. Melbourne, I., Nicol, M.: A vector-valued almost sure invariance principle for hyperbolic dynamical systems. Ann. Probab. **37**(2), 478–505 (2009)
24. Nandori, P., Szasz, D., Varju, T.: Tail asymptotics of free path lengths for the periodic Lorentz process: on Dettmann's geometric conjectures. Commun. Math. Phys. **331**, 111–137 (2014)
25. Papanicolaou, G.C.: Asymptotic analysis of transport processes. Bull. Am. Math. Soc. **81**, 330–392 (1975)
26. Spohn, H.: The Lorentz process converges to a random flight process. Commun. Math. Phys. **60**, 277–290 (1978)
27. Stromberg, K.: Probabilities on a compact group. Trans. Am. Math. Soc. **94**, 295–309 (1960)
28. Strömbergsson, A.: On the probability of a random lattice avoiding a large convex set. Proc. Lond. Math. Soc. Third Ser. **103**(6), 950–1006 (2011)
29. Szász, D., Varjú, T.: Limit laws and recurrence for the planar Lorentz process with infinite horizon. J. Stat. Phys. **129**(1), 59–80 (2007)