INVARIANCE PRINCIPLE FOR THE PERIODIC LORENTZ GAS IN THE BOLTZMANN-GRAD LIMIT

JENS MARKLOF AND BÁLINT TÓTH

ABSTRACT. In earlier work we showed that the particle displacement for the multidimensional periodic Lorentz gas, in the limit of low scatterer density (Boltzmann-Grad limit), satisfies a central limit theorem with superdiffusive scaling. The present paper extends this result to a functional central limit theorem, i.e., the weak convergence of the particle trajectory to Brownian motion.

1. INTRODUCTION

Since its inception by Lorentz in 1904, the Lorentz gas has been one of the central models of non-equilibrium statistical mechanics. It describes a gas of non-interacting point particles that move in a fixed array of spherical scatterers. If the scatterers are located on a periodic grid, we have the *periodic Lorentz gas* — the subject of this paper. An important question is whether the particle dynamics converges, in the limit of long times and a suitable rescaling of space and time units, to Brownian motion. This phenomenon is now well understood for the two-dimensional periodic Lorentz gas thanks to work of Bunimovich and Sinai [5], Melbourne and Nicols [14] (for the finite-horizon periodic Lorentz gas), Bleher [4], Szász and Varjú [16], Dolgopyat and Chernov [8] (for the infinite-horizon periodic Lorentz gas). There are however no complete proofs in higher dimensions due to the exponential growth of the complexity of singularities (Bálint and Tóth [1, 2], Chernov [6]) and, in the case of infinite horizon, the intricate geometry of channels (Dettmann [7], Nándori, Szász and Varjú [15]).

We will show here that if one considers the limit of small scatterers, then convergence to Brownian motion can indeed be established in any space dimension with a superdiffusive $\sqrt{t \log t}$ scaling of length units, where t is time measured in units of the mean collision time. This extends our earlier proof of a central limit theorem for the particle displacement [13]. The setting is as in [13]: 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. In the classic Lorentz gas the scattering is by specular reflection, but as in [11, 13] we permit here more general scattering maps corresponding to spherically symmetric potentials; see Section 2 for the precise conditions. 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 $\overline{\xi} = 1/\overline{\sigma}$ as $r \to 0$, where $\overline{\sigma} = \pi^{\frac{d-1}{2}}/\Gamma(\frac{d+1}{2})$ denotes the volume of the unit ball in $\mathbb{R}^{d-1}$.

For the classical Lorentz gas with hard sphere scatterers, the position of our test particle at time t is denoted by

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

Date: 11 June 2015.

2010 *Mathematics Subject Classification.* 37D50, 60F17, 82C40.

The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement n. 291147. The research of B.T. is partially supported by the Hungarian National Science Foundation (OTKA) through grant K100473 and by the Leverhulme Trust through International Network Grant "Laplacians, Random Walks, Quantum Spin Systems." Both authors thank the Isaac Newton Institute, Cambridge for its support and hospitality during the programmes "Periodic and Ergodic Spectral Problems" and "Random Geometry."

where $\boldsymbol{x}_0$ and $\boldsymbol{v}_0$ are 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 $\boldsymbol{x} \in \partial\mathcal{K}_r$ we choose the *outgoing* velocity $\boldsymbol{v}$, 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 $\boldsymbol{x}_t \equiv \boldsymbol{x}_0$ for all initial conditions $\boldsymbol{x}_0 \notin \mathcal{K}_r$. In the case of potentials the phase space can be similarly extended to $\mathrm{T}^1(\mathbb{R}^d)$.

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 $\mathrm{T}^1(\mathbb{R}^d)$ which we assume to be absolutely continuous with respect to the Lebesgue measure. The following theorem shows that, for random initial data and with superdiffusive normalisation, the random curve

$$[0,1] \to \mathbb{R}^d, \qquad t \mapsto \boldsymbol{X}_{T,r}(t) := \frac{\boldsymbol{x}_{tT} - \boldsymbol{x}_0}{\Sigma_d \sqrt{T \log T}}, \tag{1.2}$$

with

$$\Sigma_d^2 := \frac{2^{1-d}\overline{\sigma}}{d^2(d+1)\zeta(d)}, \tag{1.3}$$

converges in distribution in $\mathrm{C}_0([0,1])$ to the standard Brownian motion $t \mapsto \boldsymbol{W}(t)$ in $\mathbb{R}^d$ with unit covariance matrix I_d. Here $\zeta(d) := \sum_{n=1}^\infty n^{-d}$ denotes the Riemann zeta function, and $\mathrm{C}_0([0,1])$ is the space of continuous curves $[0,1] \to \mathbb{R}^d$ starting at the origin.

Theorem 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*

$$\boldsymbol{X}_{T,r} \Rightarrow \boldsymbol{W}. \tag{1.4}$$

Theorem 1 follows from its discrete-time analogue, Theorem 2 below. Denote by $\boldsymbol{q}_{n,r} = \boldsymbol{q}_{n,r}(\boldsymbol{q}_0, \boldsymbol{v}_0) \in \partial\mathcal{K}_r$ $(n = 1, 2, 3, \ldots)$ the location where the test particle with initial condition $(\boldsymbol{q}_0, \boldsymbol{v}_0)$ leaves the nth scatterer. We assume $\boldsymbol{q}_0 \in \partial\mathcal{K}_r$ is on the boundary of a scatterer and $\boldsymbol{v}_0 \in \mathrm{S}_1^{d-1}$ is an outgoing velocity, i.e., pointing away from the scatterer. By the translational invariance of the lattice, we may in fact assume without loss of generality $\boldsymbol{q}_0 \in r\mathrm{S}_1^{d-1}$. For given exit velocity $\boldsymbol{v}_0$, we write

$$\boldsymbol{q}_0 = r(\boldsymbol{s}_0 + \boldsymbol{v}_0\sqrt{1 - \|\boldsymbol{s}_0\|^2}) \tag{1.5}$$

and assume in the following that the random variable $\boldsymbol{s}_0$ is uniformly distributed in the unit disc orthogonal to $\boldsymbol{v}_0$.

By linearly interpolating between the position variables $\boldsymbol{q}_0, \boldsymbol{q}_1, \ldots, \boldsymbol{q}_n$, we obtain the piecewise linear curve

$$[0,1] \to \mathbb{R}^d, \qquad t \mapsto \boldsymbol{q}_n(t) := \boldsymbol{q}_{\lfloor nt \rfloor} + \{nt\}\left(\boldsymbol{q}_{\lfloor nt \rfloor + 1} - \boldsymbol{q}_{\lfloor nt \rfloor}\right), \tag{1.6}$$

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

$$\boldsymbol{Y}_{n,r}(t) := \frac{\boldsymbol{q}_n(t) - \boldsymbol{q}_0}{\sigma_d \sqrt{n \log n}} \tag{1.7}$$

with

$$\sigma_d^2 := \frac{2^{1-d}}{d^2(d+1)\zeta(d)}. \tag{1.8}$$

The following theorem is then the discrete-time analogue of Theorem 2.

Theorem 2. *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*

$$\boldsymbol{Y}_{n,r} \Rightarrow \boldsymbol{W}. \tag{1.9}$$

The convergence in the first limit $r \to 0$ (Boltzmann-Grad limit) is given by [11, Theorem 1.1]. That is, under the conditions of Theorem 2, for $r \to 0$ and arbitrary fixed n,

$$\boldsymbol{Y}_{n,r} \Rightarrow \boldsymbol{Y}_n. \tag{1.10}$$

We describe the limit process $t \to \boldsymbol{Y}_n(t)$ in Section 3. Section 4 shows that the $\boldsymbol{Y}_n$ converges to $\boldsymbol{W}$ in finite dimensional distribution. The last missing ingredient in the proof of Theorem 2 is then the tightness of the probability measures associated with the sequence of random curves $(\boldsymbol{Y}_n)_{n=1}^{\infty}$ in $\mathrm{C}_0([0,1])$, which is established in Section 5. Theorem 1 follows from Theorem 2 via estimates presented in Section 6.

2. The scattering map

We now specify the conditions on the scattering map that are assumed in Theorems 1 and 2. These are the same as in [13]. (The hypotheses in [11] are slightly more general.) 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

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

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. Since we assume that the scattering map is spherically symmetric, it is sufficent 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

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

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

We assume in the statements of Theorems 1 and 2 that one of the following hypotheses is true:

(A) $\theta \in \mathrm{C}^1([0,1))$ *is strictly decreasing with* $\theta(0) = \pi$ *and* $\theta(w) > 0$;
(B) $\theta \in \mathrm{C}^1([0,1))$ *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. [11]. 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

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

we have (by spherical symmetry)

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

with the matrix

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

where

$$w := \|\boldsymbol{w}\| > 0, \qquad \widehat{\boldsymbol{w}} := w^{-1}\boldsymbol{w} \in \mathrm{S}_1^{d-1}, \qquad E(\boldsymbol{x}) := \exp\begin{pmatrix} 0 & -{}^{\mathrm{t}}\boldsymbol{x} \\ \boldsymbol{x} & 0_{d-1} \end{pmatrix} \in \mathrm{SO}(d). \tag{2.8}$$

More explicitly,

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

We extend the definition of $S(\boldsymbol{w})$ to $\boldsymbol{w} = \boldsymbol{0}$ by setting $S(\boldsymbol{0}) := -I_d \in \mathrm{SO}(d)$ for d even and $S(\boldsymbol{0}) := \begin{pmatrix} -I_{d-1} & \\ & 1 \end{pmatrix} \in \mathrm{SO}(d)$ for d odd. This choice ensures that $S(\boldsymbol{0})\boldsymbol{e}_1 = -\boldsymbol{e}_1$.

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

$$\boldsymbol{v}_- = R(\boldsymbol{v}_-)\begin{pmatrix}1\\ \boldsymbol{0}\end{pmatrix}, \qquad \boldsymbol{b} = R(\boldsymbol{v}_-)\begin{pmatrix}0\\ \boldsymbol{w}\end{pmatrix}. \tag{2.10}$$

Then

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

We use an inductive argument [13, Lemma 2.1] to work out the velocity $\boldsymbol{v}_n$ after the nth collision, as well as the impact and exit parameters $\boldsymbol{b}_n$ and $\boldsymbol{s}_n$ of the nth collision: Fix $\boldsymbol{v}_0$ and $R_0 \in \mathrm{SO}(d)$ so that $\boldsymbol{v}_0 = R_0\boldsymbol{e}_1$, and denote by $(\boldsymbol{v}_n)_{n\in\mathbb{N}}$, $(\boldsymbol{b}_n)_{n\in\mathbb{N}}$, $(\boldsymbol{s}_n)_{n\in\mathbb{N}}$ the sequence of velocities, impact and exit parameters of a given particle trajectory. Then there is a *unique* sequence $(\boldsymbol{w}_n)_{n\in\mathbb{N}}$ in $\mathcal{B}_1^{d-1}$ such that for all $n \in \mathbb{N}$

$$\boldsymbol{v}_n = R_n\begin{pmatrix}1\\ \boldsymbol{0}\end{pmatrix}, \qquad \boldsymbol{b}_n = R_{n-1}\begin{pmatrix}0\\ \boldsymbol{w}_n\end{pmatrix}, \qquad \boldsymbol{s}_n = R_n\begin{pmatrix}0\\ \boldsymbol{w}_n\end{pmatrix}, \tag{2.12}$$

where

$$R_n := R_0 S(\boldsymbol{w}_1)\cdots S(\boldsymbol{w}_n). \tag{2.13}$$

3. The Boltzmann-Grad limit

To describe the limit process $\boldsymbol{Y}_n$ in (1.10), define the Markov chain (we follow here the notation of [13])

$$n \mapsto (\xi_n, \boldsymbol{\eta}_n) \tag{3.1}$$

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

$$\mathbf{P}\left((\xi_n, \boldsymbol{\eta}_n) \in \mathcal{A} \mid \xi_{n-1}, \boldsymbol{\eta}_{n-1}\right) = \int_{\mathcal{A}} \Psi_0(\boldsymbol{\eta}_{n-1}, x, \boldsymbol{z})\, dx\, d\boldsymbol{z}. \tag{3.2}$$

The transition kernel $\Psi_0(\boldsymbol{w}, x, \boldsymbol{z})$ is defined for $\boldsymbol{w}, \boldsymbol{z} \in \mathcal{B}_1^{d-1}$, $x \in \mathbb{R}_{>0}$, and is discussed in detail in [13, Section 5]. For our purposes it is sufficient to record that the kernel is independent of ξ_{n-1} and symmetric in the impact variables, i.e. $\Psi_0(\boldsymbol{w}, x, \boldsymbol{z}) = \Psi_0(\boldsymbol{z}, x, \boldsymbol{w})$. It is also independent of the choice of θ, $\mathcal{L}$ and Λ [10]. Let

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

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

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

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

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 $\Psi(x, \boldsymbol{z})$ (for the continuous time setting) or $\Psi_0(x, \boldsymbol{z})$ (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 [11, Section 6] for more details.

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

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

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:

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

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

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

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

$$\boldsymbol{Q}_n(t) := \boldsymbol{Q}_{\lfloor nt \rfloor} + \{nt\}\, \xi_{\lfloor nt \rfloor + 1} \boldsymbol{V}_{\lfloor nt \rfloor} \qquad \text{(linear interpolation);} \tag{3.12}$$

$$\boldsymbol{X}(t) := \boldsymbol{Q}_{\nu_t} + (t - \tau_{\nu_t})\boldsymbol{V}_{\nu_t} \qquad \text{(continuous time displacement).} \tag{3.13}$$

The rescaled discrete-time limiting process in (1.10) is defined by

$$\boldsymbol{Y}_n(t) := \frac{\boldsymbol{Q}_n(t)}{\sigma_d \sqrt{n \log n}}. \tag{3.14}$$

We will assume from now on that $(\xi_1, \boldsymbol{\eta}_1)$ has density $\Psi_0(x, \boldsymbol{z})$; by the arguments of [13, Section 10] this is without loss of generality. We note that for $\boldsymbol{\eta}_0$ uniformly distributed in $\mathcal{B}_1^{d-1}$,

$$\Psi_0(x, \boldsymbol{z}) = \mathbf{E}\, \Psi_0(\boldsymbol{\eta}_0, x, \boldsymbol{z}), \tag{3.15}$$

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

$$\big((\xi_n, \boldsymbol{\eta}_n)\big)_{n=0}^{\infty} \tag{3.16}$$

with the same transition probability (3.2), $\boldsymbol{\eta}_0$ uniformly distributed in $\mathcal{B}_1^{d-1}$ and ξ_0 arbitrary. Since the right hand side of (3.2) does not depend on ξ_{n-1}, the sequence

$$\underline{\boldsymbol{\eta}} = \big(\boldsymbol{\eta}_n\big)_{n=0}^{\infty}, \tag{3.17}$$

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

$$\mathbf{P}\left(\boldsymbol{\eta}_n \in \mathcal{A} \;\middle|\; \boldsymbol{\eta}_{n-1}\right) = \int_{\mathcal{A}} K_0(\boldsymbol{\eta}_{n-1}, \boldsymbol{z})\, d\boldsymbol{z} \tag{3.18}$$

where

$$K_0(\boldsymbol{w}, \boldsymbol{z}) := \int_0^\infty \Psi_0(\boldsymbol{w}, x, \boldsymbol{z}) dx. \tag{3.19}$$

The stationary measure for this Markov chain is the normalized uniform measure on $\mathcal{B}_1^{d-1}$.

We note that, if we condition on the sequence $\underline{\boldsymbol{\eta}}$, then the $\boldsymbol{V}_n$ are deterministic, and $(\xi_n)_{n=1}^{\infty}$ is a sequence of independent (but not identically distributed) random variables with distribution

$$\mathbf{P}\left(\xi_n \in (x, x+dx) \mid \underline{\boldsymbol{\eta}}\right) = \frac{\Psi_0(\boldsymbol{\eta}_{n-1}, x, \boldsymbol{\eta}_n)\, dx}{K_0(\boldsymbol{\eta}_{n-1}, \boldsymbol{\eta}_n)}. \tag{3.20}$$

4. Convergence of finite-dimensional distributions

The convergence of finite-dimensional distribution follows from the convergence of the one-dimensional distribution proved in [13] without significant further input. We include the extension argument for completeness.

Proposition 3. *Let $d \geq 2$, $\boldsymbol{v}_0 \in \mathrm{S}_1^{d-1}$ 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 $n \to \infty$,*

$$\big(\boldsymbol{Y}_n(t_1), \ldots, \boldsymbol{Y}_n(t_k)\big) \Rightarrow \big(\boldsymbol{W}(t_1), \ldots, \boldsymbol{W}(t_k)\big). \tag{4.1}$$

Proof. We may assume $t_0 := 0 < t_1 < t_2 < \ldots < t_k \leq 1$. The weak convergence (4.1) is equivalent to

$$\begin{aligned} &\big(\boldsymbol{Y}_n(t_1) - \boldsymbol{Y}_n(t_0), \ldots, \boldsymbol{Y}_n(t_k) - \boldsymbol{Y}_n(t_{k-1})\big) \\ &\qquad\qquad \Rightarrow \big(\boldsymbol{W}(t_1) - \boldsymbol{W}(t_0), \ldots, \boldsymbol{W}(t_k) - \boldsymbol{W}(t_{k-1})\big). \end{aligned} \tag{4.2}$$

Define the kd-dimensional vector

$$\boldsymbol{U}_{j,n} := \begin{pmatrix} \mathbb{1}_{\{j \leq \lfloor t_1 n \rfloor\}} \boldsymbol{V}_j \\ \mathbb{1}_{\{\lfloor t_1 n \rfloor < j \leq \lfloor t_2 n \rfloor\}} \boldsymbol{V}_j \\ \vdots \\ \mathbb{1}_{\{\lfloor t_{k-1} n \rfloor < j \leq \lfloor t_k n \rfloor\}} \boldsymbol{V}_j \end{pmatrix} \tag{4.3}$$

and

$$\boldsymbol{R}_n := \sum_{j=1}^{n} \xi_j \boldsymbol{U}_j = \begin{pmatrix} \boldsymbol{Q}_n(t_1) - \boldsymbol{Q}_n(t_0) \\ \vdots \\ \boldsymbol{Q}_n(t_k) - \boldsymbol{Q}_n(t_{k-1}) \end{pmatrix}, \qquad n \in \mathbb{N}. \tag{4.4}$$

We thus need to show that

$$\frac{\boldsymbol{R}_n}{\sigma_d \sqrt{n \log n}} \Rightarrow \begin{pmatrix} \boldsymbol{W}_n(t_1) - \boldsymbol{W}_n(t_0) \\ \vdots \\ \boldsymbol{W}_n(t_k) - \boldsymbol{W}_n(t_{k-1}) \end{pmatrix}. \tag{4.5}$$

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

$$\boldsymbol{R}_n' := \sum_{j=1}^{n} \xi_j' \boldsymbol{U}_{j-1} \tag{4.6}$$

with

$$\xi_j' := \xi_j \mathbb{1}_{\{\xi_j^2 \leq j (\log j)^{\gamma}\}} \tag{4.7}$$

for some fixed $\gamma \in (1,2)$. The following lemma tells us that it is sufficient to prove Proposition 3 for $\boldsymbol{R}_n'$ instead of $\boldsymbol{R}_n$.

Lemma 4. *We have*

$$\sup_{n \in \mathbb{N}} \|\boldsymbol{R}_n - \boldsymbol{R}_n'\| < \infty \tag{4.8}$$

almost surely.

This statement is an immediate consequence of [13, Lemma 4.1], where the bound is established for each component. The argument uses the Borel-Cantelli Lemma, which shows that $\xi'_j \neq \xi_j$ occurs for finitely many j almost surely.

To prove the central limit theorem for $\boldsymbol{R}'_n$, we center ξ'_j by setting

$$\tilde{\xi}_j = \xi'_j - m_j, \tag{4.9}$$

with the conditional expectation

$$m_j := \mathbf{E}\left(\xi'_j \mid \underline{\boldsymbol{\eta}}\right) = \frac{K_{1,r_j}(\boldsymbol{\eta}_{j-1}, \boldsymbol{\eta}_j)}{K_0(\boldsymbol{\eta}_{j-1}, \boldsymbol{\eta}_j)} \tag{4.10}$$

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

$$K_{1,r}(\boldsymbol{w}, \boldsymbol{z}) := \int_0^r x\Psi_0(\boldsymbol{w}, x, \boldsymbol{z})dx. \tag{4.11}$$

Let

$$\widetilde{\boldsymbol{R}}_n := \sum_{j=1}^n \tilde{\xi}_j \boldsymbol{U}_{j-1}. \tag{4.12}$$

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

Lemma 5. *The sequence of random variables*

$$\frac{\boldsymbol{R}'_n - \widetilde{\boldsymbol{R}}_n}{\sqrt{n \log\log n}} \tag{4.13}$$

is tight if $d = 2$, and

$$\frac{\boldsymbol{R}'_n - \widetilde{\boldsymbol{R}}_n}{\sqrt{n}} \tag{4.14}$$

is tight if $d \geq 3$.

This lemma follows directly from [13, Lemma 4.2]. The latter exploits an upper bound on the diffusive order of fluctuations of ergodic averages for Markov chains, under appropriate spectral conditions.

Lemmas 4 and 5 imply that it is sufficient to prove Proposition 3 for $\widetilde{\boldsymbol{R}}_n$ in place of $\boldsymbol{R}_n$. This will be achieved by applying the Lindeberg central limit theorem to the conditional sum as aluded to above. We begin by estimating the conditional variance. Set

$$a_j^2 := \mathbf{Var}\left(\tilde{\xi}_j \mid \underline{\boldsymbol{\eta}}\right) = \frac{K_{2,r_j}(\boldsymbol{\eta}_{j-1}, \boldsymbol{\eta}_j)}{K_0(\boldsymbol{\eta}_{j-1}, \boldsymbol{\eta}_j)} - m_j^2, \tag{4.15}$$

with

$$K_{2,r}(\boldsymbol{w}, \boldsymbol{z}) := \int_0^r x^2\Psi_0(\boldsymbol{w}, x, \boldsymbol{z})dx. \tag{4.16}$$

Lemma 6. *For $n \to \infty$,*

$$\begin{aligned} \frac{\mathbf{E}\left(\widetilde{\boldsymbol{R}}_n \otimes \widetilde{\boldsymbol{R}}_n \mid \underline{\boldsymbol{\eta}}\right)}{n \log n} &= \frac{\sum_{j=1}^n a_j^2 \boldsymbol{U}_{j-1} \otimes \boldsymbol{U}_{j-1}}{n \log n} \\ &\xrightarrow{\mathbf{P}} \sigma_d^2 \begin{pmatrix} t_1 I_d & & & \\ & (t_2 - t_1) I_d & & \\ & & \ddots & \\ & & & (t_k - t_{k-1}) I_d \end{pmatrix}. \end{aligned} \tag{4.17}$$

This follows from [13, Lemma 4.3] by observing that the different d-dimensional components of $\widetilde{\boldsymbol{R}}_n$ are independent when conditioned on $\underline{\boldsymbol{\eta}}$ and $\boldsymbol{v}_0$. The proof of [13, Lemma 4.3] establishes a weak law of large numbers for the sequence of conditional variances. This essentially relies on Chebyshev inequalities applied carefully to the truncated variables.

By taking the trace in (4.17), we have in particular (cf. [13, (4.19)])

$$\frac{A_n^2}{n\log n} \xrightarrow{\mathbf{P}} d\,\sigma_d^2 \tag{4.18}$$

for

$$A_n^2 := \sum_{j=1}^n a_j^2 = \mathbf{E}\left(\|\widetilde{\boldsymbol{Q}}_n\|^2 \mid \underline{\boldsymbol{\eta}}\right) = \mathbf{E}\left(\|\widetilde{\boldsymbol{R}}_n\|^2 \mid \underline{\boldsymbol{\eta}}\right). \tag{4.19}$$

The next lemma [13, Lemma 4.4] verifies the Lindeberg conditions for random $\underline{\boldsymbol{\eta}}$. This is the main probabilistic ingredient in the proof of the central limit theorem in [13] and relies on delicate asymptotic estimates.

Lemma 7. *For any fixed $\varepsilon > 0$,*

$$A_n^{-2}\sum_{j=1}^n \mathbf{E}\big(\tilde{\xi}_j^2 \mathbb{1}_{\{\tilde{\xi}_j^2>\varepsilon^2 A_n^2\}} \mid \underline{\boldsymbol{\eta}}\big) \xrightarrow{\mathbf{P}} 0 \tag{4.20}$$

as $n\to\infty$.

Given these lemmas, let us now conclude the proof of (4.5). The sequence of random vectors

$$\boldsymbol{Z}_n := \frac{\widetilde{\boldsymbol{R}}_n}{\sigma_d\sqrt{n\log n}} \tag{4.21}$$

is tight in $\mathbb{R}^{kd}$ because each component is tight in $\mathbb{R}^d$; the latter was proved in [13, end of Section 4]. By the Helly-Prokhorov theorem, there is an infinite subset $S_1\subset\mathbb{N}$ so that $\boldsymbol{Z}_n$ converges in distribution along $n\in S_1$ to some limit $\boldsymbol{Z}$. Assume for a contradiction that $\boldsymbol{Z}$ is *not* the right hand side of (4.5). The Borel-Cantelli lemma implies that there is an infinite subset $S_2\subset S_1$, so that in the statements of Lemmas 6 and 7 we have almost-sure convergence along $n\in S_2$:

$$\mathbf{E}\left(\boldsymbol{Z}_n\otimes\boldsymbol{Z}_n \mid \underline{\boldsymbol{\eta}}\right) \xrightarrow{\text{a.s.}} \begin{pmatrix} t_1 I_d & & & \\ & (t_2-t_1)I_d & & \\ & & \ddots & \\ & & & (t_k-t_{k-1})I_d \end{pmatrix}, \tag{4.22}$$

$$\frac{A_n^2}{n\log n} \xrightarrow{\text{a.s.}} d\,\sigma_d^2, \tag{4.23}$$

and

$$A_n^{-2}\sum_{j=1}^n \mathbf{E}\big(\tilde{\xi}_j^2 \mathbb{1}_{\{\tilde{\xi}_j^2>\varepsilon^2 A_n^2\}} \mid \underline{\boldsymbol{\eta}}\big) \xrightarrow{\text{a.s.}} 0. \tag{4.24}$$

The hypotheses of the Lindeberg central limit theorem are met, and we infer that $\boldsymbol{Z}_n$ converges to the right hand side of (4.5) for $n\to\infty$ along 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 hypothesis, and hence the right hand side of (4.5) is indeed the unique limit distribution of any weakly converging subsequence. This in turn implies that every sequence converges weakly, and therefore completes the proof of (4.5). In view of Lemmas 4 and 5, this implies Proposition 3. □

5. Tightness

We will now establish tightness for the sequence of processes $(\boldsymbol{Y}_n)_{n=1}^{\infty}$. Define

$$\xi_{j,n} := \xi_j \mathbb{1}_{\{\xi_j \le r_n\}}, \qquad r_n = \sqrt{n(\log n)^\gamma}, \tag{5.1}$$

$$\tilde{\xi}_{j,n} = \xi_{j,n} - m_{j,n}, \tag{5.2}$$

with the conditional expectation

$$m_{j,n} := \mathbf{E}\left(\xi_{j,n} \mid \underline{\boldsymbol{\eta}}\right) = \frac{K_{1,r_n}(\boldsymbol{\eta}_{j-1}, \boldsymbol{\eta}_j)}{K_0(\boldsymbol{\eta}_{j-1}, \boldsymbol{\eta}_j)}. \tag{5.3}$$

Furthermore, let

$$\boldsymbol{Q}_n^* := \sum_{j=1}^{n} \xi_{j,n} \boldsymbol{V}_{j-1}, \tag{5.4}$$

$$a_{j,n}^2 := \mathbf{Var}\left(\xi_{j,n} \mid \underline{\boldsymbol{\eta}}\right) = \frac{K_{2,r_n}(\boldsymbol{\eta}_{j-1}, \boldsymbol{\eta}_j)}{K_0(\boldsymbol{\eta}_{j-1}, \boldsymbol{\eta}_j)} - m_{j,n}^2, \tag{5.5}$$

and

$$\mathcal{A}_n^2 := \sum_{j=1}^{n} a_{j,n}^2 = \mathbf{E}\left(\|\boldsymbol{Q}_n^*\|^2 \mid \underline{\boldsymbol{\eta}}\right). \tag{5.6}$$

We split the process $\boldsymbol{Y}_n$ into four parts,

$$\boldsymbol{Y}_n = \widehat{\boldsymbol{Y}}_n + \widetilde{\boldsymbol{Y}}_n + \overline{\boldsymbol{Y}}_n + \check{\boldsymbol{Y}}_n, \tag{5.7}$$

where

$$\widehat{\boldsymbol{Y}}_n(t) = \frac{1}{\sigma_d \sqrt{n \log n}} \sum_{j=1}^{\lfloor nt \rfloor} \xi_j \mathbb{1}_{\{\xi_j > r_n\}} \boldsymbol{V}_{j-1}, \tag{5.8}$$

$$\widetilde{\boldsymbol{Y}}_n(t) = \frac{1}{\sigma_d \sqrt{n \log n}} \sum_{j=1}^{\lfloor nt \rfloor} \tilde{\xi}_{j,n} \boldsymbol{V}_{j-1}, \tag{5.9}$$

$$\overline{\boldsymbol{Y}}_n(t) = \frac{1}{\sigma_d \sqrt{n \log n}} \sum_{j=1}^{\lfloor nt \rfloor} m_{j,n} \boldsymbol{V}_{j-1}, \tag{5.10}$$

and

$$\check{\boldsymbol{Y}}_n(t) = \frac{1}{\sigma_d \sqrt{n \log n}} \{nt\}\, \xi_{\lfloor nt \rfloor + 1} \boldsymbol{V}_{\lfloor nt \rfloor}. \tag{5.11}$$

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 8. *There is $C < \infty$ such that*

$$\sup_{t \in [0,1]} \|\widehat{\boldsymbol{Y}}_n(t)\| \le \frac{C}{\sqrt{n \log n}}. \tag{5.12}$$

for all $n \in \mathbb{N}$ almost surely.

The proof of Proposition 8 is identical to that of [13, Lemma 4.1]. The key ingredient is the Borel-Cantelli Lemma, cf. the comment after Lemma 4.

Proposition 9. *There is $C < \infty$ such that, for any $\beta > 0$ and any $n \ge 2$,*

$$\mathbf{P}\left(\sup_{t \in [0,1]} \|\check{\boldsymbol{Y}}_n(t)\| \ge \beta\right) \le \frac{C}{\beta^2 \log n}. \tag{5.13}$$

Proof. We have

$$\mathbf{P}\left(\sup_{t\in[0,1]}\|\check{\boldsymbol{Y}}_n(t)\|\geq\beta\right)\leq\mathbf{P}\left(\max_{1\leq j\leq n+1}|\xi_j|\geq\beta\sigma_d\sqrt{n\log n}\right)$$
$$\leq\sum_{j=1}^{n+1}\mathbf{P}\left(|\xi_j|\geq\beta\sigma_d\sqrt{n\log n}\right)\tag{5.14}$$

where, by the asymptotic tail for the free path length distribution [12, Theorem 1.14], we have for all $j\geq 1$

$$\mathbf{P}\left(|\xi_j|\geq\beta\sigma_d\sqrt{n\log n}\right)=O\Big(\frac{1}{\beta^2 n\log n}\Big).\tag{5.15}$$

□

The estimation of $\overline{\boldsymbol{Y}}_n(t)$ relies on the following maximal inequality for martingales with stationary increments, cf. Gordin and Lifsic [9]. We denote by $\mathrm{L}^2_0(\mathcal{V},\mu)$ the orthogonal complement of the constant functions in $\mathrm{L}^2(\mathcal{V},\mu)$.

Proposition 10. *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* $\mathcal{P}f(\boldsymbol{z})=\mathbf{E}\left(f(\boldsymbol{\alpha}_n)\;\middle|\;\boldsymbol{\alpha}_{n-1}=\boldsymbol{z}\right)$*. Then, for any* $f\in\mathrm{Ran}(\mathcal{I}-\mathcal{P})$*,* $n\in\mathbb{N}$*,* $\kappa>0$*,*

$$\mathbf{P}\left(\max_{1\leq m\leq n}\left|\sum_{j=1}^{m}f(\boldsymbol{\alpha}_j)\right|\geq\kappa\sqrt{n}\right)\leq\frac{9}{\kappa^2}\left(\|g\|^2+\frac{1}{n}\|\mathcal{P}g\|^2\right),\tag{5.16}$$

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

Proof. We have

$$\sum_{j=1}^{m}f(\boldsymbol{\alpha}_j)=\sum_{j=1}^{m}\big(g(\boldsymbol{\alpha}_j)-\mathcal{P}g(\boldsymbol{\alpha}_j)\big)=\mathcal{M}_m+\mathcal{P}g(\boldsymbol{\alpha}_0)-\mathcal{P}g(\boldsymbol{\alpha}_m).\tag{5.17}$$

where

$$\mathcal{M}_m:=\sum_{j=1}^{m}\big(g(\boldsymbol{\alpha}_j)-\mathcal{P}g(\boldsymbol{\alpha}_{j-1})\big)\tag{5.18}$$

is a martingale with stationary and ergodic increments. The left hand side of (5.16) is estimated by the sum of the following three terms. The first is bounded by Doob's inequality for non-negative sub-martingales,

$$\mathbf{P}\left(\max_{1\leq m\leq n}\left|\mathcal{M}_m\right|\geq\frac{\kappa\sqrt{n}}{3}\right)\leq\frac{9}{\kappa^2 n}\mathbf{E}\left(\left|\mathcal{M}_n\right|^2\right)=\frac{9}{\kappa^2}\left(\|g\|^2-\|\mathcal{P}g\|^2\right).\tag{5.19}$$

The second follows from Chebyshev's inequality,

$$\mathbf{P}\left(\left|\mathcal{P}g(\boldsymbol{\alpha}_0)\right|\geq\frac{\kappa\sqrt{n}}{3}\right)\leq\frac{9}{\kappa^2 n}\|\mathcal{P}g\|^2,\tag{5.20}$$

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

$$\mathbf{P}\left(\max_{1\leq m\leq n}\left|\mathcal{P}g(\boldsymbol{\alpha}_m)\right|\geq\frac{\kappa\sqrt{n}}{3}\right)\leq\sum_{m=1}^{n}\mathbf{P}\left(\left|\mathcal{P}g(\boldsymbol{\alpha}_m)\right|\geq\frac{\kappa\sqrt{n}}{3}\right)\leq\frac{9}{\kappa^2}\|\mathcal{P}g\|^2.\tag{5.21}$$

□

Proposition 11. *There is* $C<\infty$ *such that, for any* $\beta>0$ *and any* $n\geq 2$*,*

$$\mathbf{P}\left(\sup_{t\in[0,1]}\|\overline{\boldsymbol{Y}}_n(t)\|\geq\beta\right)\leq\begin{cases}\frac{C\log\log n}{\beta^2\log n} & (d=2)\\ \frac{C}{\beta^2\log n} & (d\geq 3).\end{cases}\tag{5.22}$$

Proof. The plan is to apply Proposition 10 to a Markov chain defined on the state space of three consecutive velocities,

$$\mathcal{V} := \{(\boldsymbol{v}_{n-1}, \boldsymbol{v}_n, \boldsymbol{v}_{n+1}) \in (\mathrm{S}_1^{d-1})^3 : \varphi(\boldsymbol{v}_{n-1}, \boldsymbol{v}_n) > B_\theta,\ \varphi(\boldsymbol{v}_n, \boldsymbol{v}_{n+1}) > B_\theta\}, \tag{5.23}$$

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

$$p_0(\boldsymbol{v}_{n-1}, \boldsymbol{v}_n, \xi, \boldsymbol{v}_{n+1}) = \Psi_0\big(\boldsymbol{w}(\boldsymbol{v}_{n-1}, \boldsymbol{v}_n), \xi, \boldsymbol{w}'(\boldsymbol{v}_n, \boldsymbol{v}_{n+1})\big)\, \sigma(\boldsymbol{v}_n, \boldsymbol{v}_{n+1}) \tag{5.24}$$

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

$$\begin{pmatrix} 0 \\ \boldsymbol{w}(\boldsymbol{v}_{n-1}, \boldsymbol{v}_n) \end{pmatrix} = R(\boldsymbol{v}_n)^{-1} \boldsymbol{s}(\boldsymbol{v}_{n-1}, \boldsymbol{v}_n), \tag{5.25}$$

$$\begin{pmatrix} 0 \\ \boldsymbol{w}'(\boldsymbol{v}_n, \boldsymbol{v}_{n+1}) \end{pmatrix} = R(\boldsymbol{v}_n)^{-1} \boldsymbol{b}(\boldsymbol{v}_n, \boldsymbol{v}_{n+1}); \tag{5.26}$$

here $\boldsymbol{s}(\boldsymbol{v}_{n-1}, \boldsymbol{v}_n)$, $\boldsymbol{b}(\boldsymbol{v}_n, \boldsymbol{v}_{n+1}) \in \mathbb{R}^d$ are the exit and impact parameters, respectively, expressed as functions of incoming and outgoing velocities. 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 [11]. We integrate out the flight time and obtain

$$\int_0^\infty p_0(\boldsymbol{v}_{n-1}, \boldsymbol{v}_n, \xi, \boldsymbol{v}_{n+1})\, d\xi = L_0(\boldsymbol{v}_{n-1}, \boldsymbol{v}_n, \boldsymbol{v}_{n+1})\, \sigma(\boldsymbol{v}_n, \boldsymbol{v}_{n+1}) \tag{5.27}$$

with

$$L_0(\boldsymbol{v}_{n-1}, \boldsymbol{v}_n, \boldsymbol{v}_{n+1}) := K_0\big(\boldsymbol{w}(\boldsymbol{v}_{n-1}, \boldsymbol{v}_n), \boldsymbol{w}'(\boldsymbol{v}_n, \boldsymbol{v}_{n+1})\big). \tag{5.28}$$

Thus

$$n \mapsto \boldsymbol{\alpha}_n = (\boldsymbol{V}_{n-1}, \boldsymbol{V}_n, \boldsymbol{V}_{n+1}) \tag{5.29}$$

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

$$d\mu((\boldsymbol{v}_{n-1}, \boldsymbol{v}_n, \boldsymbol{v}_{n+1})) := \overline{s}^{-1} \overline{\sigma}^{-2}\, \sigma(\boldsymbol{v}_{n-1}, \boldsymbol{v}_n)\, \sigma(\boldsymbol{v}_n, \boldsymbol{v}_{n+1})\, d\boldsymbol{v}_{n-1}\, d\boldsymbol{v}_n\, d\boldsymbol{v}_{n+1} \tag{5.30}$$

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

$$\mathbf{P}\big(\boldsymbol{\alpha}_n \in \mathcal{A} \,\big|\, \boldsymbol{\alpha}_{n-1} = \boldsymbol{z}\big) = \int_{\mathcal{A}} \mathcal{K}(\boldsymbol{z}, \boldsymbol{z}')\, d\mu(\boldsymbol{z}') \tag{5.31}$$

with kernel

$$\begin{aligned} &\mathcal{K}\big((\boldsymbol{v}_{n-2}, \boldsymbol{v}_{n-1}, \boldsymbol{v}_n), (\boldsymbol{v}'_{n-1}, \boldsymbol{v}'_n, \boldsymbol{v}'_{n+1})\big) \\ &\qquad = \overline{s}\, \overline{\sigma}^2\, \frac{\delta(\boldsymbol{v}_{n-1}, \boldsymbol{v}'_{n-1})\, \delta(\boldsymbol{v}_n, \boldsymbol{v}'_n)\, L_0(\boldsymbol{v}'_{n-1}, \boldsymbol{v}'_n, \boldsymbol{v}'_{n+1})}{\sigma(\boldsymbol{v}_{n-1}, \boldsymbol{v}_n)}. \end{aligned} \tag{5.32}$$

Note that the kernel $\mathcal{K}(\boldsymbol{z}, \boldsymbol{z}')$ 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

$$\begin{aligned} \mathcal{P} f(\boldsymbol{z}) &= \mathbf{E}\left(f(\boldsymbol{\alpha}_n) \,\big|\, \boldsymbol{\alpha}_{n-1} = \boldsymbol{z}\right) \\ &= \int \mathcal{K}(\boldsymbol{z}, \boldsymbol{z}') f(\boldsymbol{z}')\, d\mu(\boldsymbol{z}'). \end{aligned} \tag{5.33}$$

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

$$\begin{aligned} \mathcal{P}^m f(\boldsymbol{z}) &= \mathbf{E}\left(f(\boldsymbol{\alpha}_n) \,\big|\, \boldsymbol{\alpha}_{n-1} = \boldsymbol{z}\right) \\ &= \int \mathcal{K}^{(m)}(\boldsymbol{z}, \boldsymbol{z}') f(\boldsymbol{z}')\, d\mu(\boldsymbol{z}'). \end{aligned} \tag{5.34}$$

The second power reads

$$\mathcal{K}^{(2)}\big((\boldsymbol{v}_{n-2},\boldsymbol{v}_{n-1},\boldsymbol{v}_n),(\boldsymbol{v}'_n,\boldsymbol{v}'_{n+1},\boldsymbol{v}'_{n+2})\big) = \overline{s}\,\overline{\sigma}^2 L_0(\boldsymbol{v}_{n-1},\boldsymbol{v}_n,\boldsymbol{v}'_{n+1})\, L_0(\boldsymbol{v}'_n,\boldsymbol{v}'_{n+1},\boldsymbol{v}'_{n+2})\,\delta(\boldsymbol{v}_n,\boldsymbol{v}'_n). \tag{5.35}$$

As to the third power,

$$\mathcal{K}^{(3)}\big((\boldsymbol{v}_{n-2},\boldsymbol{v}_{n-1},\boldsymbol{v}_n),(\boldsymbol{v}'_{n+1},\boldsymbol{v}'_{n+2},\boldsymbol{v}'_{n+3})\big) = \overline{s}\,\overline{\sigma}^2 L_0(\boldsymbol{v}_{n-1},\boldsymbol{v}_n,\boldsymbol{v}'_{n+1})\, L_0(\boldsymbol{v}_n,\boldsymbol{v}'_{n+1},\boldsymbol{v}'_{n+2})\, L_0(\boldsymbol{v}'_{n+1},\boldsymbol{v}'_{n+2},\boldsymbol{v}'_{n+3})\,\sigma(\boldsymbol{v}_n,\boldsymbol{v}'_{n+1}). \tag{5.36}$$

In view of the lower bound on $K_0(\boldsymbol{w},\boldsymbol{z})$ [13, Lemma 6.1/6.2], we have

$$\mathcal{K}^{(3)}\big((\boldsymbol{v}_{n-2},\boldsymbol{v}_{n-1},\boldsymbol{v}_n),(\boldsymbol{v}'_{n+1},\boldsymbol{v}'_{n+2},\boldsymbol{v}'_{n+3})\big) \geq \frac{\overline{s}\,\overline{\sigma}^2}{(2^d\overline{\sigma}\zeta(d))^3}\,\sigma(\boldsymbol{v}_n,\boldsymbol{v}'_{n+1}). \tag{5.37}$$

The fourth power reads

$$\begin{aligned}\mathcal{K}^{(4)}\big((\boldsymbol{v}_{n-2},\boldsymbol{v}_{n-1},\boldsymbol{v}_n),(\boldsymbol{v}'_{n+2},\boldsymbol{v}'_{n+3},\boldsymbol{v}'_{n+4})\big) &= \overline{s}\,\overline{\sigma}^2 \int_{\mathrm{S}_1^{d-1}} L_0(\boldsymbol{v}_{n-1},\boldsymbol{v}_n,\tilde{\boldsymbol{v}}_{n+1})\, L_0(\boldsymbol{v}_n,\tilde{\boldsymbol{v}}_{n+1},\boldsymbol{v}'_{n+2})\, L_0(\tilde{\boldsymbol{v}}_{n+1},\boldsymbol{v}'_{n+2},\boldsymbol{v}'_{n+3}) \\ &\quad\times L_0(\boldsymbol{v}'_{n+2},\boldsymbol{v}'_{n+3},\boldsymbol{v}'_{n+4})\,\sigma(\boldsymbol{v}_n,\tilde{\boldsymbol{v}}_{n+1})\,\sigma(\tilde{\boldsymbol{v}}_{n+1},\boldsymbol{v}'_{n+2})\,d\tilde{\boldsymbol{v}}_{n+1}.\end{aligned} \tag{5.38}$$

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

$$\mathcal{K}^{(4)}\big((\boldsymbol{v}_{n-2},\boldsymbol{v}_{n-1},\boldsymbol{v}_n),(\boldsymbol{v}'_{n+2},\boldsymbol{v}'_{n+3},\boldsymbol{v}'_{n+4})\big) \geq \frac{\overline{s}\,\overline{\sigma}^2}{(2^d\overline{\sigma}\zeta(d))^4}\int_{\mathrm{S}_1^{d-1}} \sigma(\boldsymbol{v}_n,\tilde{\boldsymbol{v}}_{n+1})\,\sigma(\tilde{\boldsymbol{v}}_{n+1},\boldsymbol{v}'_{n+2})\,d\tilde{\boldsymbol{v}}_{n+1}. \tag{5.39}$$

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

$$\begin{aligned}\mathcal{K}^{(m)}\big((\boldsymbol{v}_{n-2},\boldsymbol{v}_{n-1},\boldsymbol{v}_n),(\boldsymbol{v}'_{n+m-2},\boldsymbol{v}'_{n+m-1},\boldsymbol{v}'_{n+m})\big) &\geq \frac{\overline{s}\,\overline{\sigma}^2}{(2^d\overline{\sigma}\zeta(d))^m}\int_{\mathrm{S}_1^{d-1}} \sigma(\boldsymbol{v}_n,\tilde{\boldsymbol{v}}_{n+1})\,\sigma(\tilde{\boldsymbol{v}}_{n+1},\tilde{\boldsymbol{v}}_{n+2})\cdots \\ &\quad\cdots\sigma(\tilde{\boldsymbol{v}}_{n+m-4},\tilde{\boldsymbol{v}}_{n+m-3})\,\sigma(\tilde{\boldsymbol{v}}_{n+m-3},\boldsymbol{v}'_{n+m-2})\,d\tilde{\boldsymbol{v}}_{n+1}\cdots d\tilde{\boldsymbol{v}}_{n+m-3}.\end{aligned} \tag{5.40}$$

For the class of scattering maps considered here, there exists a finite m such that the integral on the right-hand side has a uniform positive lower bound for all $\boldsymbol{v}_n,\boldsymbol{v}'_{n+m-2}\in\mathrm{S}_1^{d-1}$. Hence

$$\inf_{\boldsymbol{z},\boldsymbol{z}'\in\mathcal{V}}\mathcal{K}^{(m)}(\boldsymbol{z},\boldsymbol{z}')>0. \tag{5.41}$$

This, by the classic Doeblin argument (see e.g. [13, Section 7]), implies that $\mathcal{P}^m$, and therefore $\mathcal{P}$, has a spectral gap.

Define

$$L_{1,r}(\boldsymbol{v}_{n-1},\boldsymbol{v}_n,\boldsymbol{v}_{n+1}) := K_{1,r}\big(\boldsymbol{w}(\boldsymbol{v}_{n-1},\boldsymbol{v}_n),\boldsymbol{w}'(\boldsymbol{v}_n,\boldsymbol{v}_{n+1})\big), \tag{5.42}$$

and the function $\ell_r:\mathcal{V}\to\mathbb{R}_{\geq 0}$ by

$$\ell_r((\boldsymbol{v}_{n-1},\boldsymbol{v}_n,\boldsymbol{v}_{n+1})) := \frac{L_{1,r}(\boldsymbol{v}_{n-1},\boldsymbol{v}_n,\boldsymbol{v}_{n+1})}{L_0(\boldsymbol{v}_{n-1},\boldsymbol{v}_n,\boldsymbol{v}_{n+1})}. \tag{5.43}$$

We then recover the random variable (5.3) via

$$m_{j,n}=\ell_{r_n}(\boldsymbol{\alpha}_n)=\ell_{r_n}((\boldsymbol{V}_{n-1},\boldsymbol{V}_n,\boldsymbol{V}_{n+1})). \tag{5.44}$$

To conclude the proof, apply Proposition 10 with the (n-dependent) choice of $f\in\mathrm{L}^2(\mathcal{V},d\mu)$,

$$f((\boldsymbol{v}_{n-1},\boldsymbol{v}_n,\boldsymbol{v}_{n+1})) = \boldsymbol{e}\cdot\boldsymbol{v}_n\,\ell_{r_n}((\boldsymbol{v}_{n-1},\boldsymbol{v}_n,\boldsymbol{v}_{n+1})) \tag{5.45}$$

with an arbitrary fixed $\boldsymbol{e} \in \mathrm{S}_1^{d-1}$. 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 $R \in \mathrm{SO}(d)$, we have

$$\int_{\mathcal{V}} f(\boldsymbol{z})\, d\mu(\boldsymbol{z}) = 0, \tag{5.46}$$

i.e., $f \in \mathrm{L}_0^2(\mathcal{V}, \mu)$. Thanks to the spectral gap of $\mathcal{P}$, there is a constant $M < \infty$, such that for all $f \in \mathrm{L}_0^2(\mathcal{V}, \mu)$,

$$\|(\mathcal{I} - \mathcal{P})^{-1} f\|^2 \le M \|f\|^2 \tag{5.47}$$

Finally, the estimate for $\mathbf{E}(m_n^2)$ in [13, 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 11 thus follows from Proposition 10 with $\kappa = \beta\sqrt{\log n}$. □

Propositions 8, 9 and 11 establish that the tightness for $(\boldsymbol{Y}_n)_{n=1}^\infty$ is implied by the tightness for $(\widetilde{\boldsymbol{Y}}_n)_{n=1}^\infty$. To prove the latter, we use the following classical characterization of tightness.

Theorem 12. [3, Theorem 8.3] *The sequence $(\lambda_n)_{n=1}^\infty$ of probability measures in $\mathrm{C}_0([0,1])$ 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*

$$\frac{1}{\delta} \sup_{t \in [0,1]} \lambda_n \left(\sup_{s \in [t, t+\delta] \cap [0,1]} \left\| \boldsymbol{X}(s) - \boldsymbol{X}(t) \right\| \ge \beta \right) < \epsilon. \tag{5.48}$$

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

Lemma 13. [3, 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 $\boldsymbol{S}_m = \boldsymbol{\xi}_1 + \ldots + \boldsymbol{\xi}_m$ and $s_m^2 = \sigma_1^2 + \ldots + \sigma_m^2$. Then*

$$\mathbf{P}\left(\max_{i \le m} \|\boldsymbol{S}_i\| \ge \lambda s_m \right) \le 2\, \mathbf{P}\left(\|\boldsymbol{S}_m\| \ge (\lambda - \sqrt{2}) s_m \right). \tag{5.49}$$

The following proposition verifies the hypothesis of Theorem 12, and thus proves that the sequence of probability measures corresponding to $(\widetilde{\boldsymbol{Y}}_n)_{n=1}^\infty$ is tight.

Proposition 14. *For every $\epsilon > 0$, $\beta > 0$ there exist $\delta < 1$, $n_0 < \infty$ such that for all $n \ge n_0$ we have*

$$\frac{1}{\delta} \sup_{t \in [0,1]} \mathbf{P}\left(\sup_{s \in [t, t+\delta] \cap [0,1]} \left\| \widetilde{\boldsymbol{Y}}_n(s) - \widetilde{\boldsymbol{Y}}_n(t) \right\| \ge \beta \right) < \epsilon. \tag{5.50}$$

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

$$\frac{1}{\delta} \mathbf{P}\left(\max_{nt < m \le n(t+\delta)} \left\| \sum_{j=\lfloor nt \rfloor + 1}^{m} \tilde{\xi}_{j,n} \boldsymbol{V}_{j-1} \right\| \ge \beta\sqrt{n \log n} \right) < \epsilon. \tag{5.51}$$

To prove this fact, note first that

$$\begin{aligned} \mathbf{P}\left(\max_{nt < m \le n(t+\delta)} \left\| \sum_{j=\lfloor nt \rfloor + 1}^{m} \tilde{\xi}_{j,n} \boldsymbol{V}_{j-1} \right\| \ge \beta\sqrt{n \log n} \right) \\ = \mathbf{E}\left(\mathbf{P}\left(\max_{nt < m \le n(t+\delta)} \left\| \sum_{j=\lfloor nt \rfloor + 1}^{m} \tilde{\xi}_{j,n} \boldsymbol{V}_{j-1} \right\| \ge \beta\sqrt{n \log n} \;\middle|\; \underline{\boldsymbol{\eta}} \right) \right). \end{aligned} \tag{5.52}$$

The maximal inequality (5.49) yields

$$\mathbf{P}\left(\max_{nt<m\le n(t+\delta)} \Big\| \sum_{j=\lfloor nt\rfloor+1}^{m} \tilde{\xi}_{j,n} \boldsymbol{V}_{j-1} \Big\| \ge \beta\sqrt{n\log n} \;\Big|\; \underline{\boldsymbol{\eta}} \right) \le 2\,\mathbf{P}\left(\Big\| \sum_{j=\lfloor nt\rfloor+1}^{\lfloor n(t+\delta)\rfloor} \tilde{\xi}_{j,n} \boldsymbol{V}_{j-1} \Big\| \ge \beta\sqrt{n\log n} - \sqrt{2(\mathcal{A}^2_{\lfloor n(t+\delta)\rfloor} - \mathcal{A}^2_{\lfloor nt\rfloor})} \;\Big|\; \underline{\boldsymbol{\eta}} \right), \tag{5.53}$$

which in turn implies, after taking expectation values,

$$\mathbf{P}\left(\max_{nt<m\le n(t+\delta)} \Big\| \sum_{j=\lfloor nt\rfloor+1}^{m} \tilde{\xi}_{j,n} \boldsymbol{V}_{j-1} \Big\| \ge \beta\sqrt{n\log n} \right) \le 2\,\mathbf{P}\left(\Big\| \sum_{j=\lfloor nt\rfloor+1}^{\lfloor n(t+\delta)\rfloor} \tilde{\xi}_{j,n} \boldsymbol{V}_{j-1} \Big\| \ge \beta\sqrt{n\log n} - \sqrt{2(\mathcal{A}^2_{\lfloor n(t+\delta)\rfloor} - \mathcal{A}^2_{\lfloor nt\rfloor})} \right). \tag{5.54}$$

The latter is bounded above by

$$\le 2\,\mathbf{P}\left(\Big\| \sum_{j=\lfloor nt\rfloor+1}^{\lfloor n(t+\delta)\rfloor} \tilde{\xi}_{j,n} \boldsymbol{V}_{j-1} \Big\| \ge \big(\beta - K\sqrt{\delta}\big)\sqrt{n\log n} \right) + 2\,\mathbf{P}\left(\mathcal{A}^2_{\lfloor n(t+\delta)\rfloor} - \mathcal{A}^2_{\lfloor nt\rfloor} > K^2\delta n\log n \right), \tag{5.55}$$

for any constant $K \ge 0$. By adapting the proof of [13, Lemma 4.3] (replacing $j(\log j)^\gamma$ by $n(\log n)^\gamma$), one can prove that the analogue of (4.18) reads

$$\frac{\mathcal{A}_n^2}{n\log n} \xrightarrow{\mathbf{P}} d\,\sigma_d^2. \tag{5.56}$$

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

$$\lim_{n\to\infty} \mathbf{P}\left(\mathcal{A}^2_{\lfloor n(t+\delta)\rfloor} - \mathcal{A}^2_{\lfloor nt\rfloor} > K^2\delta\, n\log n \right) = 0 \tag{5.57}$$

uniformly in t, δ, which takes care of the second term in (5.55). The first term in (5.55) is estimated by the central limit theorem of our previous paper [13, 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 $n \ge n_0$

$$\mathbf{P}\left(\Big\| \sum_{j=\lfloor nt\rfloor+1}^{\lfloor n(t+\delta)\rfloor} \tilde{\xi}_{j,n} \boldsymbol{V}_{j-1} \Big\| \ge \big(\beta - K\sqrt{\delta}\big)\sqrt{n\log n} \right) \le \frac{1}{(2\pi)^{d/2}} \int_{\|\boldsymbol{x}\| > \frac{\beta - K\sqrt{\delta}}{\sigma_d\sqrt{\delta}}} \mathrm{e}^{-\frac12\|\boldsymbol{x}\|^2} d\boldsymbol{x} + \epsilon\,\delta = O(\epsilon\,\delta). \tag{5.58}$$

Note that here we have applied [13, Theorem 3.2 (ii)] to the truncated $\tilde{\xi}_{j,n}$ rather than ξ_j, which is justified by the analogue of [13, Lemmas 4.1, 4.2] with $j(\log j)^\gamma$ replaced by $n(\log n)^\gamma$. This completes the proof. □

6. Theorem 2 implies Theorem 1

We now turn to the continuous time process. Boltzmann-Grad limit $r \to 0$ is covered by [11, Theorem 1.2], which tells us that for arbitrary fixed T,

$$\boldsymbol{X}_{T,r} \Rightarrow \boldsymbol{X}_T \tag{6.1}$$

where

$$\boldsymbol{X}_T(t) = \frac{\boldsymbol{X}(tT)}{\Sigma_d\sqrt{T\log T}} = \frac{\boldsymbol{Q}_{\nu_{tT}} + (t - \tau_{\nu_{tT}})\boldsymbol{V}_{\nu_{tT}}}{\Sigma_d\sqrt{T\log T}}; \tag{6.2}$$

recall (3.8)–(3.13).

The convergence of finite-dimensional distributions of $\boldsymbol{X}_T$ to $\boldsymbol{W}$ follows (within the framework of Section 4) from the same estimates as in [13, Section 11]. What remains is to show tightness in $\mathrm{C}([0,1])$ for the family of processes $(\boldsymbol{X}_T)_{T\geq 1}$. By a simple scaling argument, it is sufficient to prove tightness for the sequence $(\boldsymbol{X}_n)_{n\in\mathbb{N}}$.

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

$$T(\theta) := \tau_{\lfloor\theta\rfloor} + \{\theta\}\xi_{\lfloor\theta\rfloor+1}, \qquad \Theta(t) := \nu_t + \frac{t-\tau_{\nu_t}}{\xi_{\nu_t+1}}. \tag{6.3}$$

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 $t \to \boldsymbol{X}_n(t)$ and $\theta \to \boldsymbol{Q}_n(\theta)$ by this time change: We have $\boldsymbol{X}(t) = \boldsymbol{Q}_1(\Theta(t))$ and therefore

$$\boldsymbol{X}(nt) = \boldsymbol{Q}_1(\Theta(nt)) = \boldsymbol{Q}_n(n^{-1}\Theta(nt)) = \boldsymbol{Q}_n(\Theta_n(t)), \tag{6.4}$$

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

$$\boldsymbol{X}_n(t) = \overline{\xi}^{1/2}\, \boldsymbol{Y}_n(\Theta_n(t)), \tag{6.5}$$

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

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

$$Z_n(\theta) := \frac{T(n\theta) - n\theta\overline{\xi}}{\sigma_d\sqrt{d\, n\log n}}. \tag{6.6}$$

The following lemma says that Z_n converges to one-dimensional Brownian motion W; cf. also [13, Lemma 11.3].

Lemma 15. *For $n \to \infty$,*

$$Z_n \Rightarrow W. \tag{6.7}$$

The proof of this lemma is a simpler variant of the already established weak convergence $\boldsymbol{Y}_n \Rightarrow \boldsymbol{W}$.

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

The modulus of continuity of a curve $\boldsymbol{X} \in \mathrm{C}_0([0,b])$ is defined as

$$\omega_{\boldsymbol{X}}^{[0,b]}(\delta) := \sup_{\substack{0\leq s,t\leq b\\ |t-s|\leq\delta}} \left\|\boldsymbol{X}(s) - \boldsymbol{X}(t)\right\|. \tag{6.8}$$

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

Lemma 16. *For every $\beta > 0$, $\epsilon > 0$ there exist $\delta < 1$ and $n_0 < \infty$ such that for all $n \geq n_0$*

$$\mathbf{P}\left(\omega_{\boldsymbol{X}_n}^{[0,1]}(\delta) > \beta\right) < \epsilon. \tag{6.9}$$

Proof. Notice that

$$\mathbf{P}\left(\omega_{\boldsymbol{X}_n}^{[0,1]}(\delta) > \beta\right) \leq \mathbf{P}\left(\omega_{\overline{\xi}^{1/2}\boldsymbol{Y}_n}^{[0,2\overline{\sigma}]} \circ \omega_{\Theta_n}^{[0,1]}(\delta) > \beta\right) + \mathbf{P}\left(\Theta_n(1+\delta) > 2\overline{\sigma}\right). \tag{6.10}$$

For $n \to \infty$, we have $\Theta_n(1+\delta) \xrightarrow{\mathbf{P}} (1+\delta)\overline{\sigma} < 2\overline{\sigma}$ (see the remark after Lemma 15), and therefore

$$\lim_{n\to\infty} \mathbf{P}\left(\Theta_n(1+\delta) > 2\overline{\sigma}\right) = 0. \tag{6.11}$$

The claim now follows from the tightness of $(\boldsymbol{Y}_n)_{n=1}^{\infty}$ established in Theorem 2, and from the tightness of $(\Theta_n)_{n=1}^{\infty}$ which follows from the remark after Lemma 15. □

This concludes the proof of Theorem 1.

References

1. P Bálint and I P Tóth, *Exponential decay of correlations in multi-dimensional dispersing billiards*, Annales Henri Poincaré **9** (2008), no. 7, 1309–1369.
2. ———, *Example for exponential growth of complexity in a finite horizon multi-dimensional dispersing billiard*, Nonlinearity **25** (2012), no. 5, 1275–1297.
3. P Billingsley, Convergence of probability measures. First edition. John Wiley & Sons, Inc., New York, 1968. (Note that the second edition is not compatible with the references made.)
4. P M Bleher, *Statistical properties of two-dimensional periodic Lorentz gas with infinite horizon*, Journal of Statistical Physics **66** (1992), no. 1-2, 315–373.
5. L A Bunimovich and Ya G Sinai, *Statistical properties of Lorentz gas with periodic configuration of scatterers*, Communications in Mathematical Physics **78** (1980), no. 4, 479–497.
6. N I Chernov, *Statistical properties of the periodic Lorentz gas. Multidimensional case*, Journal of Statistical Physics **74** (1994), no. 1-2, 11–53.
7. C P Dettmann, *New horizons in multidimensional diffusion: the Lorentz gas and the Riemann hypothesis*, Journal of Statistical Physics **146** (2012), no. 1, 181–204.
8. D I Dolgopyat and N I Chernov, *Anomalous current in periodic Lorentz gases with an infinite horizon*, Rossii skaya Akademiya Nauk. Moskovskoe Matematicheskoe Obshchestvo. Uspekhi Matematicheskikh Nauk **64** (2009), no. 4(388), 73–124.
9. M I Gordin and B A Lifsic, *Central limit theorem for stationary Markov processes*, Dokl. Akad. Nauk SSSR **239** (1978), no. 4, 766–767.
10. J Marklof and A Strömbergsson, *The distribution of free path lengths in the periodic Lorentz gas and related lattice point problems*, Annals of Mathematics. Second Series **172** (2010), no. 3, 1949–2033.
11. ———, *The Boltzmann-Grad limit of the periodic Lorentz gas*, Annals of Mathematics. Second Series **174** (2011), no. 1, 225–298.
12. ———, *The periodic Lorentz gas in the Boltzmann-Grad limit: asymptotic estimates*, Geometric and Functional Analysis **21** (2011), no. 3, 560–647.
13. J Marklof and B Tóth, *Superdiffusion in the periodic Lorentz gas*, arXiv:1403.6024.
14. I Melbourne and M Nicol, *A vector-valued almost sure invariance principle for hyperbolic dynamical systems*, The Annals of Probability **37** (2009), no. 2, 478–505.
15. P Nándori, D Szász, and T Varjú, *Tail asymptotics of free path lengths for the periodic Lorentz process. On Dettmann's geometric conjectures*, Comm. Math. Phys. **331** (2014), no. 1, 111–137.
16. D Szász and T Varjú, *Limit laws and recurrence for the planar Lorentz process with infinite horizon*, J. Stat. Phys. **129** (2007), no. 1, 59–80.

Jens Marklof, School of Mathematics, University of Bristol, Bristol BS8 1TW, U.K.
j.marklof@bristol.ac.uk

Bálint Tóth, School of Mathematics, University of Bristol, Bristol BS8 1TW, U.K.; MTA-BME Stochastics Research Group, Budapest, Hungary; Rényi Institute, Budapest, Hungary
balint.toth@bristol.ac.uk, balint@math.bme.hu