*The Annals of Applied Probability*
2017, Vol. 27, No. 1, 65–90
DOI: 10.1214/16-AAP1197


# POLYNOMIAL CONVERGENCE TO EQUILIBRIUM FOR A SYSTEM OF INTERACTING PARTICLES

BY YAO LI[*,†] AND LAI-SANG YOUNG[†,1]

*University of Massachusetts Amherst[*] and New York University[†]*

We consider a stochastic particle system in which a finite number of particles interact with one another via a common energy tank. Interaction rate for each particle is proportional to the square root of its kinetic energy, as is consistent with analogous mechanical models. Our main result is that the rate of convergence to equilibrium for such a system is $\sim t^{-2}$, more precisely it is faster than a constant times $t^{-2+\varepsilon}$ for any $\varepsilon > 0$. A discussion of exponential vs. polynomial convergence for similar particle systems is included.

This paper is about dynamical models of (large numbers of) interacting particles, a topic of fundamental importance in both dynamical systems and statistical mechanics. Our focus is on the speed of convergence to equilibrium, equivalently the rate of decay of time correlations. On a fixed energy surface, Liouville measure, which describes the states of a system in equilibrium, does not depend on the dynamics generated by the Hamiltonian, but once the system is taken *out of equilibrium*, the speed with which it returns to equilibrium can be affected by dynamical details. One of the purposes of this paper is to call attention to the fact that for particle systems, this convergence can be fast or slow depending on how the particles interact.

While Hamiltonian models are considered to be physically more realistic than stochastic ones, questions of ergodicity and mixing for general Hamiltonian systems are out of reach at the present time, let alone the rate of mixing. Simplifications on the level of modeling are necessary if one is to gain insight into the problem. Since chaotic dynamics are known to produce statistics very similar to those of genuinely random stochastic processes [2, 7, 38, 40, 47], it seems logical to first tackle stochastic models designed to capture similar underlying phenomena.

The following model of binary collisions introduced by Kac [24] half a century ago as an idealization of Boltzmann dynamics was in this spirit. In Kac's model, the velocities of $N$ particles are described (abstractly) by $N$ real numbers $v_1, v_2, \ldots, v_N$, so that the system has total energy $\sum_{i=1}^{N} v_i^2 = E$. An exponential

Received March 2015; revised March 2016.

[1]Supported in part by NSF Grant DMS-13-63161.

*MSC2010 subject classifications.* Primary 60J25, 60K35, 82C22; secondary 60J20.

*Key words and phrases.* Interacting particle model, polynomial convergence rate, Markov jump process.

clock rings with rate $N$. When it rings, a pair of particles, $i$ and $j$, is randomly chosen and assumed to interact, resulting in new velocities, $v_i'$ and $v_j'$, given by

$$v_i' = (\cos\theta)v_i - (\sin\theta)v_j,$$
$$v_j' = (\sin\theta)v_i + (\cos\theta)v_j,$$

where $\theta \in [0, 2\pi)$ is uniformly distributed. This model has been much studied. Among other things, it has been shown that its infinitesimal generator has a spectral gap uniformly bounded away from zero in size for all $N$ [4, 21, 33]. Models with energy-dependent interactions, which are more realistic than the constant rate of interaction in the original model, have also been studied [5], as have other variants of this model; see, for example, [19, 39] for binary collision processes on lattices and [18] for extensions to quantum $N$-body problems.

In general, for systems with direct particle-particle interactions and an interaction potential that falls off with distance, it is very difficult to identify a simple stochastic rule that captures faithfully the deterministic dynamics. In this paper, we consider a class of particle systems for which such modeling is more straightforward, namely when the particles do not interact with one another directly but only via their "environment," or a "hub." Concrete examples of mechanical models of this type were introduced in [34, 37] and studied later in [10–13, 25, 30, 31, 44]. In these models, the "environment" is symbolized by the kinetic energy stored in rotating disks placed at various locations in the physical domain. When a particle collides with a disk, energy is exchanged in accordance with a rule consistent with energy and angular momentum conservation; point particles do not "see" each other otherwise. See Figure 2. The models considered in the present paper are a stochastic version of these mechanical models; details are given in Sections 1.1 and 1.2.

An example of the type of stochastic modification we make is that we "forget" the precise location of a particle, and replace the time to its next collision by an exponential random variable with mean $\propto \frac{1}{\sqrt{e}}$ where $e$ is the kinetic energy of the particle. This idea was also used in [13], and is consistent with the statistics produced by chaotic dynamical systems. More detailed justification is given in Section 1.1.

We prove for our models that the speed of convergence to equilibrium is not exponential but polynomial. More precisely, we show that for any $\gamma > 0$, this rate is faster than $\sim t^{\gamma-2}$. Because the rate of interaction is $\propto\sqrt{e}$, it is not hard to see that convergence rate cannot be faster than $\sim t^{-2}$. Thus, our results are sharp, and to our knowledge they are new; a literature search has not turned up comparable results involving polynomial rates of convergence. The closest that we are aware of are [45, 46], which showed slower than exponential convergence for certain mechanical models with special properties (e.g., particles interacting only with heat baths, or particle systems on physical domains with special geometry).

The speed of convergence to equilibrium, equivalently the rate of decay of time correlations, impacts the type of probabilistic limit laws obeyed by the system. We do not pursue that here as these questions will take us too far afield, but remark only on some immediate consequences: To have a strictly convex rate function of the large deviation principle, a spectral gap is usually necessary [1, 26, 42, 43]. With polynomial rates of convergence, the large deviation rate function, if it exists, can have a flat section [27]. Also due to the polynomial ergodicity, one can only expect the Markov chains central limit theorem to hold for bounded observables; see Theorem 1.6 for the detail.

The main ideas of our proof are as follows: since low-energy particles are the source of slow convergence, we call a state of the system, equivalently an energy configuration, "active" if every particle carries an energy above a certain minimum. Starting from the set of active states we prove a Doeblin-type condition, suggesting exponential correlation decay for an induced process. We then return to the full system, and propose to view the dynamics as having been refreshed, or renewed, each time a trajectory returns to the set of active states. This puts us in a framework bearing some resemblance to renewal processes, for which it has been shown that the speed of convergence to equilibrium is determined by the moments of renewal times. Following ideas from renewal theory, we seek to control first passage times to the set of active states. This is done by constructing a suitable Lyapunov function; see Section 2.

*Polynomial vs. exponential convergence*: *further examples*. The root cause of the slow convergence in our model is that once a particle acquires a low energy in an interaction, it simply stays "frozen" until its clock rings again; there is no way to activate it sooner. This need not be the case in models with direct particle-particle interactions, if another particle can pass by and activate a slow particle. The question of exponential vs. polynomial rates of convergence to equilibrium is most transparent in the setting of *one particle per site*, *nearest-neighbor interactions*, an example of which is the locally confined disk models introduced in [3] and studied in [15, 16]: A linear chain of cells is connected by openings. Inside each cell is a single finite-size convex body (hard disk), the diameter of which exceeds that of the opening so it is trapped, but adjacent disks can meet and exchange energy; see Figure 1. For these models, the rate of convergence hinges on whether a disk can be completely out of reach of its neighbors. When the openings are large enough, heuristic argument and numerical simulations both give exponential convergence. On the other hand, if the openings between cells are small enough that a disk can get entirely out of reach of its neighbors, then a phenomenon similar to that in the present paper can occur: it is easy to *prove* that the rate of mixing cannot be faster than $t^{-2}$; see [28], which contains also a numerical study confirming that the rate of mixing is $\sim t^{-2}$, and the rate of interaction between disks with kinetic energies $e_i$ and $e_{i+1}$ can be approximated by $\sim\sqrt{\min\{e_i, e_{i+1}\}}$.

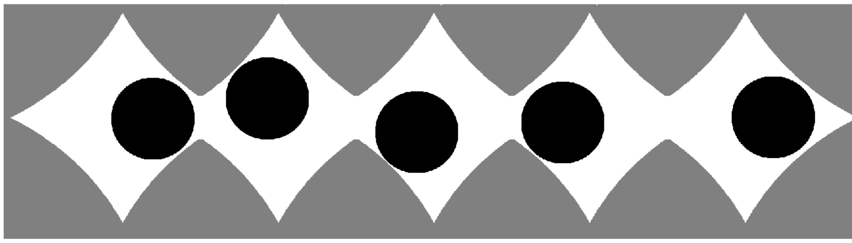

FIG. 1. *Locally confined hard disks model. Whether the system converges to equilibrium at exponential or polynomial speeds depends on its geometric configuration, specifically whether or not there are positions where a disk (black) can be out of reach of its neighbors.*

We comment on related works: In a nonrigorous derivation, Gaspard and Gilbert [17] argued for the same model that under certain assumptions, the rate of interaction between the $i$th and $(i+1)$st disks is $\sim\sqrt{e_i + e_{i+1}}$. Assuming this interaction rate, [19, 29, 39] proved exponential rates of convergence for stochastic versions of these models. To our knowledge, this interaction rate appears in a certain rare interaction limit (when the openings between cells tend to zero), and involves a rescaling of time. Without taking any limits or rescaling time, it is a simple mathematical fact that correlations in the mechanical models above cannot decay faster than $t^{-2}$ when the disks can "hide" from their neighbors.

*Organization of this paper*. Section 1 contains a precise model description and statement of results. The bulk of the technical work goes into the construction of a Lyapunov function; this is carried out in Section 2. In Section 3, we use this Lyapunov function to deduce the desired results on polynomial convergence to equilibrium.

## 1. Model and results.

As explained in the Introduction, the models considered in this paper are stochastic versions of some known mechanical models. We begin with a review of these mechanical models, followed by a discussion of the rationale for replacing the deterministic dynamics by Markovian dynamics. Section 1.2 contains the precise definitions of the models studied in the rest of this paper, and the statement of results are announced in Section 1.3.

### 1.1. *Mechanical models with particle-disk interactions*.

We review here a class of models consisting of a rotating disk and a finite number of particles in a closed domain, energy being exchanged when a particle collides with the disk. The rules of energy exchange are borrowed from [34]; see also [37]. These models, both in and out of equilibrium, were studied in [13].

A precise model description is as follows: Let $\Gamma \subset \mathbb{R}^2$ be a bounded domain with concave piecewise $C^3$ boundary; see Figure 2 for an example. In the interior of $\Gamma$ is a rotating disk $D$, nailed down at its center and rotating freely, carrying with

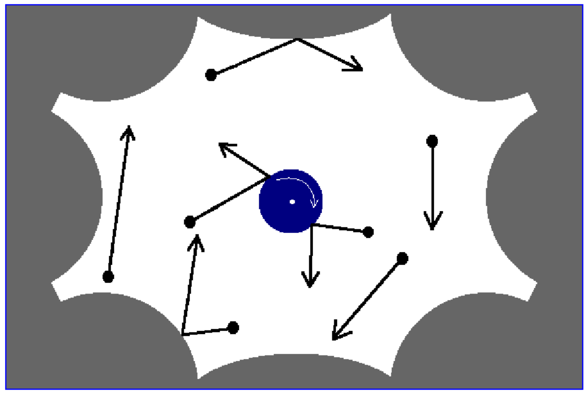

FIG. 2. *Example of a mechanical system that motivated the present study*: *Particles in a domain* $\Gamma$ (*white*) *are scattered as they are reflected off* $\partial\Gamma$, *and energy is exchanged when a particle collides with the rotating disk* (*blue*) *nailed down at the center of the domain.*

it a finite amount of kinetic energy. In the region $\Gamma \setminus D$ are $m$ point particles, each moving with uniform motion until it collides with $\partial\Gamma$ or $D$. Upon collision with $\partial\Gamma$, a particle is reflected elastically. Upon collision with $D$, energy is exchanged according to the following rule: Let $v$ be the velocity of the particle just prior to collision, $v = v_n + v_t$ its decomposition into components that are normal and tangential to the disk, and let $\omega$ denote the angular velocity of the disk. If $'$ denotes the corresponding velocities following the collision, then from the conservation of energy and angular momentum, one obtains, following [34],

$$\begin{aligned} v_n' &= -v_n, \\ v_t' &= v_t - \frac{2\eta}{1+\eta}(v_t - R\omega), \\ R\omega' &= R\omega + \frac{2}{1+\eta}(v_t - R\omega). \end{aligned}$$

In these formulas, $\bar{m}$ is the mass of the particle, $R$ is the radius of the disk, $\theta$ is the moment of inertia of the disk and $\eta = \theta/(\bar{m}R^2)$. This is a complete description of the model.

Choosing $R = \eta = 1$ leads to the especially simple equations

$$v_n' = -v_n, \qquad v_t' = \omega, \qquad \omega' = v_t. \tag{1.1}$$

For simplicity, we will work with these special parameters, though conceptually it makes no difference in the present study.

*Connection to stochastic model.* Though easy to describe, an analysis of the mechanical model above is considerably outside of the reach of current dynamical systems techniques. Thus, we seek to simplify the model while retaining its essential characteristics, including the way in which energy is transferred among particles. By "forgetting" the precise locations of particles in the cell and their directions of travel, as well as the direction of rotation of the disk, we turn the deterministic dynamical system above into a Markov process. Specifically, the times to energy exchange for a particle are determined by exponential distributions with mean $x^{-1/2}$ where $x$ is the instantaneous kinetic energy of the particle, and the repartitioning of energy at exchanges are as in (1.1) assuming random angles of incidence. Details are given in Section 1.2.

We provide below some heuristic justification for the memory loss and interaction proposed in the last paragraph.

First, we explain the rationale behind neglecting precise locations within a cell. Billiard systems on domains with concave boundaries (or scatterers) are well known to exhibit chaotic or hyperbolic behavior [6, 40]. *Hyperbolicity* here refers to exponential divergence of nearby orbits, a property that leads to rapid loss of memory of trajectory history. By taking the rotating disk in our model to be relatively small, between energy exchanges a typical particle trajectory is reflected many times as it bounces off the walls of the domain. (Adding more scatterers in $\Gamma \setminus D$ as was done in [30] will further enhance mixing.) As our system is a hyperbolic billiard between collisions with the rotating disk, the rapid loss of memory gives justification for neglecting precise locations within a cell.

Next we explain the use of exponential random variables to describe the times between collisions. Another well-known fact for strongly hyperbolic systems including billiards is that for points randomly distributed in a specific region, return times to this region have exponentially small tails [47]. Thus, for particles that emerge from an energy exchange with a fixed energy but randomly distributed otherwise in terms of location and angle, we can expect the times to their next collision with the disk to have an exponentially small tail.

Finally, fixing initial location and direction of travel, the time for a particle to reach a pre-specified region is proportional to its speed; that is, the rationale for assuming mean collision time is proportional to $x^{-1/2}$.

For another confirmation of the close connection between the stochastic model in Section 1.2 and the mechanical model above, notice that modulo constants their invariant measures coincide; see the remark following Proposition 1.1.

1.2. *Precise description of stochastic model.* The stochastic model considered in the rest of this paper is a time-homogeneous Markov jump process $\mathbf{x}_t, t \geq 0$, with

$$\mathbf{x}_t = (x_t^1, \dots, x_t^m, y_t).$$

Here, $m$ is a fixed positive integer, $x_t^1, \dots, x_t^m$ are the energies of the $m$ particles at time $t$, and $y_t$ is the energy of the disk, which we regard from here on as an abstract "energy tank." As the domain is assumed to be closed, total energy remains constant in time, that is, there exists a constant $\bar{E} > 0$ such that $\sum_i x_t^i + y_t = \bar{E}$ for all $t \geq 0$. Thus, the state space of $\mathbf{x}_t$ is the open $(m+1)$-dimensional simplex

$$\mathbf{\Delta} = \mathbf{\Delta}^{m+1}(\bar{E}) = \left\{ (x^1, \ldots, x^m, y) \in \mathbb{R}_+^{m+1} \middle| y + \sum_{i=1}^{m} x^i = \bar{E} \right\}.$$

As in the mechanical model in Section 1.1, the particles in this system do not interact directly with one another. Instead, they interact via the energy tank, which symbolizes the "environment" within the domain, and it is these particle-tank interactions that give rise to the jumps in the process. The rules of interaction are as follows: Particle $i$ carries a clock that rings at an exponential rate equal to $\sqrt{x_t^i}$; notice that this rate changes each time the particle acquires a new energy. The clocks carried by different particles are independent of one another and of history. When its clock rings, a particle exchanges energy with the tank according to the same rule used in the mechanical model: Suppose the clock of particle $i$ rings at time $t$, and let $\mathbf{x}_{t+} = (x_{t+}^1, x_{t+}^2, \ldots, x_{t+}^m, y_{t+})$ denote the state immediately following the interaction at time $t$. We assume that the directions of motion of the particles in the mechanical model are uniformly distributed as given by Liouville measure, so that the cosines of their angles of incidence with the rotating disk are uniformly distributed. The rules for updating, that is, (1.1), then translate into

$$x_{t+}^i = y_t + (1 - u^2) x_t^i, \qquad y_{t+} = u^2 x_t^i \quad \text{and} \quad x_{t+}^j = x_t^j \qquad \text{for } j \neq i, \tag{1.2}$$

where $u \in (0, 1)$ is a uniform random variable. For a detailed calculation, see [30].

The transition probabilities above together with an initial condition $\mathbf{x}_0$ defines the Markov process $\mathbf{x}_t$. The notation $\mathbf{x}_t = (x_t^1, \ldots, x_t^m, y_t)$ is used throughout; in particular, $x^i$ is used exclusively to denote the energy of the $i$th particle, not the $i$th power of $x$.

We fix also the following notation: For $t \geq 0$ and $\mathbf{x} \in \mathbf{\Delta}$, let $P^t(\mathbf{x}, \cdot)$ be the transition probabilities of the process $\mathbf{x}_t$. That is to say, $P^t(\mathbf{x}, \cdot)$ is the Borel probability distribution on $\mathbf{\Delta}$ describing the possible states of the system $t$ units of time later given that its initial condition is $\mathbf{x}$. To simplify notation, we use the same notation for the left and right operators generated by $P^t$:

$$(P^t \xi)(\mathbf{x}) = \int_{\mathbf{\Delta}} P^t(\mathbf{x}, \mathrm{d}\mathbf{y}) \xi(\mathbf{y})$$

for a measurable function $\xi$ on $\mathbf{\Delta}$, and

$$(\mu P^t)(A) = \int_{\mathbf{\Delta}} P^t(\mathbf{x}, A) \mu(\mathrm{d}\mathbf{x})$$

for a probability measure $\mu$ on $\mathbf{\Delta}$. Finally we say $\mu$ is an invariant measure for the process $\mathbf{x}_t$ if $\mu P^t = \mu$ for all $t > 0$.

1.3. *Statement of results.*

PROPOSITION 1.1. *The probability measure $\pi$ with density*

$$\rho(x^1, \dots, x^m, y) = \frac{1}{Z} y^{-1/2},$$

*where $Z$ is a normalizing constant is an invariant measure for the process $\mathbf{x}_t$.*

By the change of variables $x = |\mathbf{v}_i|^2$ and $y = \tilde{\omega}^2$, one sees that $\pi$ coincides with Liouville measure on a fixed energy shell for a Hamiltonian system with $H = |\mathbf{v}_i|^2 + \tilde{\omega}^2$. Here, $\mathbf{v}_i$ is the velocity of the $i$th particle, and $\tilde{\omega}$ is the angular velocity of the rotating disk.

THEOREM 1.2 (Uniqueness of invariant measure). *The measure $\pi$ in Proposition 1.1 is the unique invariant probability for $\mathbf{x}_t$; hence it is ergodic.*

THEOREM 1.3 (Speed of convergence to equilibrium). *For every $\mathbf{x} \in \mathbf{\Delta}$ and $\gamma > 0$,*

$$\lim_{t\to\infty} t^{2-\gamma} \| P^t(\mathbf{x}, \cdot) - \pi \|_{\mathrm{TV}} = 0,$$

*where $\| \cdot \|_{\mathrm{TV}}$ is the total variational norm.*

Theorem 1.3 is in fact deduced from Theorem 1.4 below. For $\delta > 0$, let $\mathcal{M}_\delta$ be the collection of probability measures $\mu$ on $\mathbf{\Delta}$ such that

$$\int_{\mathbf{\Delta}} \left( \sum_{k=1}^{m} (x^k)^{2\delta-1} + y^{\delta-\frac{1}{2}} \right) \mu(\mathrm{d}\mathbf{x}) < \infty.$$

THEOREM 1.4 (Polynomial contraction of Markov operator). *For any $\gamma > 0$ and $\mu, \nu \in \mathcal{M}_{\gamma/8}$,*

$$\lim_{t\to\infty} t^{2-\gamma} \| \mu P^t - \nu P^t \|_{\mathrm{TV}} = 0.$$

The following simple argument shows that the bound in Theorem 1.4 is tight: Consider, for example, two initial distributions $\mu$ and $\nu$ that differ by a positive amount when restricted to the set $B_\epsilon := \{x^i < \epsilon\}$ for some fixed $i$. For definiteness, let us assume that for all small enough $\epsilon$, $\mu|_{B_\epsilon} \le c\pi|_{B_\epsilon}$ and $\nu|_{B_\epsilon} \ge c'\pi|_{B_\epsilon}$ for some $c < 1 < c'$. Since $\pi(B_\epsilon) \propto \epsilon$, $x^i < \frac{1}{t^2}$ implies that the probability with respect to $\pi$ of the $i$th clock ringing before time $t$ is $<1 - e^{-1}$. It follows that

$$\| \mu P^t - \nu P^t \|_{\mathrm{TV}} \ge \| (\mu P^t - \nu P^t)|_{\{x^i < \frac{1}{t^2}\}} \|_{\mathrm{TV}} \ge \text{constant} \cdot \frac{1}{t^2}.$$

Another corollary of Theorem 1.4 is the rate of decay of time correlations.

THEOREM 1.5 (Polynomial correlation decay). *For any $\gamma > 0$ and $\mu \in \mathcal{M}_{\gamma/8}$, let $\xi$ and $\zeta \in L^\infty(\mathbf{\Delta})$. Then*

$$\left| \int_{\mathbf{\Delta}} (P^t \zeta)(\mathbf{x}) \xi(\mathbf{x}) \mu(\mathrm{d}\mathbf{x}) - \int_{\mathbf{\Delta}} (P^t \zeta)(\mathbf{x}) \mu(\mathrm{d}\mathbf{x}) \int_{\mathbf{\Delta}} \xi(\mathbf{x}) \mu(\mathrm{d}\mathbf{x}) \right| = o\left(\frac{1}{t^{2-\gamma}}\right)$$

*as $t \to \infty$.*

The Markov chain central limit theorem can also be implied by the $t^{-(2-\gamma)}$ rate of convergence.

THEOREM 1.6 (Markov chain central limit theorem). *Let $f : \mathbf{\Delta} \to \mathbb{R}$ be a $\pi$-almost surely uniformly bounded Borel function. Let $\{f_n^\delta\}_{n=1}^\infty = \{f(\mathbf{x}_0), f(\mathbf{x}_\delta), \dots, f(\mathbf{x}_{n\delta}), \dots\}$ be a sequence of observables, where $\delta > 0$ is a constant. Let*

$$\bar{f} = \frac{1}{n} \sum_{i=0}^{n} f_n^\delta.$$

*Then for any initial distribution $\mathbf{x}_0$, as $n \to \infty$,*

$$\sqrt{n}(\bar{f} - \mathbb{E}_\pi f) \xrightarrow{d} N(0, \sigma_f^2),$$

*where*

$$\sigma_f^2 := \mathrm{var}_\pi\{f(\mathbf{x}_0)\} + 2 \sum_{i=1}^{\infty} \mathrm{cov}\{f(\mathbf{x}_0), f(\mathbf{x}_{i\delta})\} < \infty.$$

## 2. Construction of Lyapunov function.

Let $\mathbf{\Delta}$ be as in Section 1.2. For $\alpha < \frac{1}{2}$, we define $V = V_\alpha : \mathbf{\Delta} \to \mathbb{R}^+$ by

$$V(\mathbf{x}) = V_\alpha(\mathbf{x}) = \sum_{i=1}^{m} (x^i)^{-2\alpha} + y^{-\alpha}.$$

Our main technical result is the following.

THEOREM 2.1. *For $\alpha < \frac{1}{2}$ close enough to $\frac{1}{2}$ and $h > 0$ small enough, there exist $c_0, M > 0$ depending on $\alpha$ and $h$ such that for $V = V_\alpha$ and $\beta = 1 - (4\alpha)^{-1}$,*

$$(P^h V)(\mathbf{x}) - V(\mathbf{x}) \le -c_0 V(\mathbf{x})^\beta$$

*for every $\mathbf{x} \in \{V > M\}$.*

The motivation for this choice of Lyapunov function is as follows. As noted in the Introduction, low energy particles are our main concern, for they are not expected to interact for a long time, and that slows down the mixing process. For this reason, a desirable Lyapunov function should satisfy $V(\mathbf{x}) \to \infty$ as

$\mathbf{x} \to \partial \mathbf{\Delta}$. We explain heuristically why one may expect something along the lines of $P^h V - V \sim -hV^{1/2}$, corresponding to $\alpha, \beta \approx 2$: Assume $x^1 \ll 1$ is the smallest particle energy. Then $V(\mathbf{x}) \sim (x^1)^{-1}$. If the clock of particle 1 rings on the time interval $[0, h)$ and $y$ is "large," then the expected drop of $V(\mathbf{x})$ following an interaction is $\sim (x^1)^{-1} \sim V(\mathbf{x})$. But the probability that the clock of particle 1 will ring exactly once before time $h$ is $\sim h\sqrt{x^1}$. This means the expected drop of $V(\mathbf{x})$ is $\sim h(x^1)^{-1/2} \sim hV^{1/2}$.

It is convenient to use the following equivalent description of $\Phi_t$: Starting from $t = 0$, a clock rings at time $\tau_1$ where $\tau_1$ is an exponential random variable with mean $(\sum_{i=1}^m \sqrt{x_0^i})^{-1}$. When this clock rings, energy exchange takes place between exactly one particle and the tank, and the probability that particle $i$ is chosen is

$$\frac{\sqrt{x_0^i}}{\sum_{i=1}^m \sqrt{x_0^i}}.$$

The rule of energy redistribution is determined by equation (1.2) as before, and this process is repeated, that is, at time $\tau_2$, an exponential random variable with mean $(\sum_{i=1}^m \sqrt{x_{\tau_1^+}^i})^{-1}$, the clock rings again, and so on.

We begin with the following technical estimate.

LEMMA 2.2. *There exist constants* $\epsilon_0 > 0$ *and* $c^* > 0$ *such that*

$$\mathbb{E}\big[V(\mathbf{x}_{\tau_1^+})|\mathbf{x}_0\big] \le V(\mathbf{x}_0) - \frac{c^*}{\sum_{i=1}^m \sqrt{x_0^i}} V(\mathbf{x}_0)^\beta$$

*for every* $\mathbf{x}_0 \in B$, *where*

$$B = \{\mathbf{x} \in \mathbf{\Delta} | y < \epsilon_0,\ \text{or}\ x^i < 4^{-\frac{1}{2\alpha}} \epsilon_0 \text{ for some } i \in \{1, \dots, m\}\}.$$

PROOF. By definition,

$$\mathbb{E}\big[V(\mathbf{x}_{\tau_1^+})|\mathbf{x}_0\big] = V(\mathbf{x}_0) + \frac{1}{\sum_{i=1}^m \sqrt{x_0^i}} \sum_{i=1}^m Q_i,$$

where

$$Q_i = \sqrt{x_0^i} \left\{ \int_0^1 [(x_0^i(1-u^2) + y_0)^{-2\alpha} + (x_0^i u^2)^{-\alpha}]\, du - [(x_0^i)^{-2\alpha} + y_0^{-\alpha}] \right\},$$

that is, we need to show $\sum Q_i \le -c^* V(\mathbf{x}_0)^\beta$ for some $c^* > 0$. In the rest of the proof, we will omit the subscript 0 in $\mathbf{x}_0$, $x_0^i$ and $y_0$, and write

$$C_1 = \int_0^1 (1-u^2)^{-2\alpha}\, du \quad \text{and} \quad C_2 = \int_0^1 u^{-2\alpha}\, du,$$

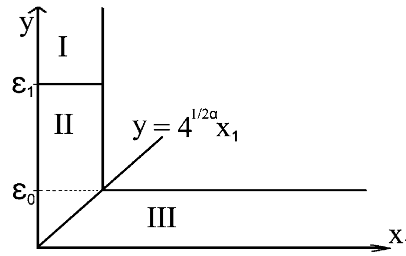


FIG. 3. *Decomposition of neighborhood of* $\partial\mathbf{\Delta}$.

noting that $C_1, C_2 < \infty$ for $\alpha < \frac{1}{2}$. We will use many times the bound

$$Q_i \le \sqrt{x^i}\{\min\{C_1(x^i)^{-2\alpha}, y^{-2\alpha}\} + C_2(x^i)^{-\alpha} - (x^i)^{-2\alpha} - y^{-\alpha}\}. \tag{2.1}$$

Without loss of generality, assume

$$x^1 = \min_{1\le i\le m} x^i.$$

Let $0 < \epsilon_0 \ll \epsilon_1 \ll \bar{E}$ be two small numbers to be determined. We decompose $B$, the neighborhood of $\partial\mathbf{\Delta}$ in the statement of the lemma, into three regions (see Figure 3) and analyze each one as follows:

REGION I. $4^{\frac{1}{2\alpha}}x^1 < \epsilon_0, y \ge \epsilon_1$.

With regard to lowering $V$, we clearly have the most to gain if particle 1 interacts with the tank: Applying (2.1) to $x^1$ and substituting in $y \ge \epsilon_1$, we obtain

$$Q_1 \le \sqrt{x^1}\cdot\{(\epsilon_1)^{-2\alpha} + C_2(x^1)^{-\alpha} - (x^1)^{-2\alpha}\}.$$

Using $4^{\frac{1}{2\alpha}}x^1 < \epsilon_0 \ll \epsilon_1$, we see that the third term dominates. Hence,

$$Q_1 \le -\frac{1}{2}(x^1)^{-2\alpha+\frac{1}{2}}.$$

For $i \ne 1$, we consider separately the following two cases: For $x^i < \frac{1}{2}\epsilon_1 < \frac{1}{2}y$, we have

$$Q_i \le \sqrt{x^i}\{(2x^i)^{-2\alpha} + C_2(x^i)^{-\alpha} - (x^i)^{-2\alpha}\}, \tag{2.2}$$

which is $<0$ since the last term dominates. If $x^i \ge \frac{1}{2}\epsilon_1$, then from (2.1) we obtain

$$Q_i \le \sqrt{x^i}\{C_1(x^i)^{-2\alpha} + C_2(x^i)^{-\alpha}\} \le C'(\epsilon_1)^{-2\alpha+\frac{1}{2}}$$

for some $C'$ independent of $\epsilon_0$ or $\epsilon_1$. Notice that we have used $\frac{1}{2} - 2\alpha < 0$, or $\alpha > \frac{1}{4}$.

Altogether, we have shown, using $\epsilon_0 \ll \epsilon_1 \ll 1$, that

$$\sum_{i=1}^{m} Q_i \le -\frac{1}{2}(x^1)^{-2\alpha+\frac{1}{2}} + (m-1)C'(\epsilon_1)^{-2\alpha+\frac{1}{2}} \le -\frac{1}{3}(x^1)^{-2\alpha+\frac{1}{2}}.$$

It follows from $V(\mathbf{x}) \le (m+1)(x^1)^{-2\alpha}$ that this is $\le -\frac{1}{3(m+1)}V(\mathbf{x})^{\beta}$.

REGION II. $4^{\frac{1}{2\alpha}}x^1 < \epsilon_0, 4^{\frac{1}{2\alpha}}x^1 < y < \epsilon_1$.

For $i = 1$, applying (2.1) and using $y > 4^{\frac{1}{2\alpha}}x^1$, we obtain

$$Q_1 \le \sqrt{x^1}\{(4^{\frac{1}{2\alpha}}x^1)^{-2\alpha} + C_2(x^1)^{-\alpha} - (x^1)^{-2\alpha}\} \le -\frac{1}{2}(x^1)^{-2\alpha+\frac{1}{2}}.$$

For $i \ne 1$, if $x^i < \frac{1}{2}y$, then the situation is as in (2.2), and $Q_i < 0$. The case where $x^i \ge \frac{1}{2}y$ is one of the more delicate: Applying (2.1), we obtain

$$Q_i \le C''x^{-2\alpha+\frac{1}{2}} - \sqrt{x}y^{-\alpha} \le C'''y^{-2\alpha+\frac{1}{2}} - \sqrt{x}y^{-\alpha}.$$

Without loss of generality, assume $x^2, \ldots, x^k \ge \frac{1}{2}y$, and $x^j < \frac{1}{2}y$ for all $j > k$. Then $\max\{x^2, \ldots, x^k\} > \frac{\bar{E}}{2m}$. Therefore,

$$\begin{aligned}\sum_{i=2}^{k} Q_i &\le (k-1)C'''y^{-2\alpha+\frac{1}{2}} - \left(\sum_{i=2}^{k}\sqrt{x^i}\right)y^{-\alpha}\\ &\le y^{-\alpha}\left[mC'''y^{-\alpha+\frac{1}{2}} - \sqrt{\frac{\bar{E}}{2m}}\right].\end{aligned}$$

As $y < \epsilon_1$ and $\alpha < \frac{1}{2}$, the quantity in square brackets is $<0$ provided $\epsilon_1$ is sufficiently small.

Thus arguing as in Region I, we have shown that

$$\sum_{i=1}^{m} Q_i \le -\frac{1}{2}(x^1)^{-2\alpha+\frac{1}{2}} < -\frac{1}{2(m+1)}V(\mathbf{x})^{\beta}.$$

REGION III. $4^{\frac{1}{2\alpha}}x^1 \ge y, y < \epsilon_0$.

Since $x^i \ge x^1 \ge 4^{-\frac{1}{2\alpha}}y$ for all $i$, a calculation analogous to that in Region II gives

$$\sum_{i=1}^{m} Q_i \le y^{-\alpha}\left[mC''''y^{-\alpha+\frac{1}{2}} - \sqrt{\frac{\bar{E}}{2m}}\right] < -\frac{1}{2}\sqrt{\frac{\bar{E}}{2m}}y^{-\alpha}$$

provided $\epsilon_0$ is small enough. Since $V(\mathbf{x}) \le m(x^1)^{-2\alpha} + y^{-\alpha} \le (4m+1)y^{-2\alpha}$, it follows that

$$\sum_{i=1}^m Q_i \le -\frac{1}{2}\sqrt{\frac{\bar{E}}{2m}} \cdot \frac{1}{\sqrt{4m+1}} V(\mathbf{x})^{\frac{1}{2}} \le -\frac{1}{2}\sqrt{\frac{\bar{E}}{2m}} \cdot \frac{1}{\sqrt{4m+1}} V(\mathbf{x})^{\beta}$$

since $\beta < \frac{1}{2}$.

The assertion is proved since it holds for $\mathbf{x}_0$ in all three regions of $B$. $\square$

PROOF OF THEOREM 2.1. Let $\tau_1 < \tau_2 < \cdots$ be the times of clock rings as defined in the paragraph preceding the statement of Lemma 2.2, and let $B$ be the neighborhood of $\partial\mathbf{\Delta}$ in Lemma 2.2. Letting $\tau_0 = 0$, we have shown that for any $n \ge 0$, if $\mathbf{x}_{\tau_n^+} \in B$, then

$$\mathbb{E}\big[V(\mathbf{x}_{\tau_{n+1}^+})|\mathbf{x}_{\tau_n^+}\big] \le V(\mathbf{x}_{\tau_n^+}) - \frac{c^*}{\sum_{i=1}^m \sqrt{x_{\tau_n^+}^i}} V(\mathbf{x}_{\tau_n^+})^{\beta}. \tag{2.3}$$

For $\mathbf{x}_{\tau_n^+} \notin B$, we will use the bound

$$\mathbb{E}\big[V(\mathbf{x}_{\tau_{n+1}^+})|\mathbf{x}_{\tau_n^+}\big] \le M_0 + M_1, \tag{2.4}$$

where

$$M_0 = \sup_{\mathbf{x}\in\mathbf{\Delta}\setminus B} V(\mathbf{x}) \quad \text{and} \quad M_1 = \sup_{\mathbf{x}\in\mathbf{\Delta}\setminus B} \frac{1}{\sum_{i=1}^m \sqrt{x^i}} \sum_{i=1}^m Q_i(\mathbf{x}).$$

It is easy to check that $M_0, M_1 < \infty$.

We now use these estimates to deduce a bound for $P^h V$ for fixed $h > 0$. Let $S = \inf\{n, \tau_n > h\}$, and define $\hat{\tau}_n = \min\{\tau_n, \tau_{S-1}\}$. Then

$$P^h V(\mathbf{x}) = \lim_{n\to\infty} \mathbb{E}\big[V(\mathbf{x}_{\hat{\tau}_n^+})\mathbf{1}_{S\le n+1}|\mathbf{x}_0 = \mathbf{x}\big] \le \lim_{n\to\infty} \mathbb{E}\big[V(\mathbf{x}_{\hat{\tau}_n^+})|\mathbf{x}_0 = \mathbf{x}\big].$$

We will prove a uniform bound for $\mathbb{E}[V(\mathbf{x}_{\hat{\tau}_n^+})|\mathbf{x}_0 = \mathbf{x}]$ for all $n \ge 1$.

First, assuming the worse of (2.3) and (2.4), we have

$$\mathbb{E}\big[V(\mathbf{x}_{\tau_{n+1}^+})|\tau_{n+1} \le h\big] \le E\big[V(\mathbf{x}_{\tau_n^+})|\tau_{n+1} \le h\big] + M_0 + M_1 \tag{2.5}$$

for every $n \ge 0$. Notice that conditioning on $\tau_{n+1} \le h$ does not affect the bounds in (2.3) and (2.4) because given $\mathbf{x}_{\tau_n^+}$, $\mathbf{x}_{\tau_{n+1}^+}$ is independent of $\tau_{n+1} - \tau_n$. Second, as $\sum_{i=1}^m \sqrt{x^i} \le \sqrt{m\bar{E}}$ for all $\mathbf{x} \in \mathbf{\Delta}$, we have, for every $\mathbf{x}_{\tau_n^+} \in \mathbf{\Delta}$,

$$\mathbb{P}[\tau_{n+1} \le h|\mathbf{x}_{\tau_n^+}, \tau_n \le h] \le \big(1 - e^{-h\sqrt{m\bar{E}}}\big)\mathbb{P}[\tau_n \le h],$$

so that inductively,

$$\mathbb{P}[\tau_{n+1} \le h|\mathbf{x}_0] \le \big(1 - e^{-h\sqrt{m\bar{E}}}\big)^{n+1} \qquad \text{for } n \ge 0. \tag{2.6}$$

The estimates (2.5) and (2.6) together imply the following: Given $\mathbf{x}_0 = \mathbf{x}$,

$$\begin{aligned}\mathbb{E}\big[V(\mathbf{x}_{\hat{\tau}^+_{n+1}})\big] &= \mathbb{E}\big[V(\mathbf{x}_{\hat{\tau}^+_{n+1}})|\tau_{n+1} > h\big]\cdot\mathbb{P}[\tau_{n+1} > h] \\ &\quad + \mathbb{E}\big[V(\mathbf{x}_{\hat{\tau}^+_{n+1}})|\tau_{n+1} \le h\big]\cdot\mathbb{P}[\tau_{n+1} \le h] \\ &\le \mathbb{E}\big[V(\mathbf{x}_{\hat{\tau}^+_n})|\tau_{n+1} > h\big]\cdot\mathbb{P}[\tau_{n+1} > h] \\ &\quad + \big(\mathbb{E}\big[V(\mathbf{x}_{\hat{\tau}^+_n})|\tau_{n+1} \le h\big] + M_0 + M_1\big)\cdot\mathbb{P}[\tau_{n+1} \le h] \\ &\le \mathbb{E}\big[V(\mathbf{x}_{\hat{\tau}^+_n})\big] + (M_0 + M_1)\big(1 - e^{-h\sqrt{m\bar{E}}}\big)^{n+1}.\end{aligned}$$

Summing over $n$, this gives

$$P^h V(\mathbf{x}) \le \mathbb{E}\big[V(\mathbf{x}_{\hat{\tau}^+_1})|\mathbf{x}_0 = \mathbf{x}\big] + \frac{M_0 + M_1}{e^{-h\sqrt{m\bar{E}}}}.$$

Let $h > 0$ be small enough so that for all $\mathbf{x} \in \boldsymbol{\Delta}$,

$$\mathbb{P}[\tau_1 \le h|\mathbf{x}_0 = \mathbf{x}] = 1 - e^{-h\sum_{i=1}^m \sqrt{x^i}} > \frac{h}{2}\sum_{i=1}^m \sqrt{x^i}.$$

This is the only condition we impose on $h$.

We choose $M'$ large enough so that $\{V > M'\} \subset B$, and consider $\mathbf{x}_0 \in \{V > M'\}$. Noting again that $\mathbb{E}[V(\mathbf{x}_{\tau_1^+})]$ is independent of $\tau_1$, we have, by Lemma 2.2,

$$\begin{aligned}\mathbb{E}\big[V(\mathbf{x}_{\hat{\tau}^+_1})\big] &= \mathbb{E}\big[V(\mathbf{x}_{\tau^+_1})|\tau_1 \le h\big]\cdot\mathbb{P}[\tau_1 \le h] + V(\mathbf{x}_0)\cdot\mathbb{P}[\tau_1 > h] \\ &\le \left(V(\mathbf{x}_0) - \frac{c^*}{\sum\sqrt{x^i}}V(\mathbf{x}_0)^\beta\right)\cdot\mathbb{P}[\tau_1 \le h] + V(\mathbf{x}_0)\cdot\mathbb{P}[\tau_1 > h] \\ &\le V(\mathbf{x}_0) - c^*\frac{h}{2}V(\mathbf{x}_0)^\beta.\end{aligned}$$

This gives

$$P^h V(\mathbf{x}) \le V(\mathbf{x}) - c^*\frac{h}{2}V(\mathbf{x})^\beta + (M_0 + M_1)e^{h\sqrt{m\bar{E}}}.$$

To complete the proof of Theorem 2.1, it suffices to replace $M'$ by a large enough number $M$ so that for $\mathbf{x} \in \{V > M\}$, the constant $(M_0 + M_1)e^{h\sqrt{m\bar{E}}}$ is absorbed into $c_0 V(\mathbf{x})^\beta$ for $c_0 = c^*\frac{h}{4}$. $\square$

We record for later use the following fact that follows from the proof above.

COROLLARY 2.3.

$$\sup_{\mathbf{x}\notin B} P^h V(\mathbf{x}) < \infty.$$

**3. Completing the proofs.** After some preliminaries in Section 3.1, we proceed to the main task of this section, the deduction of Theorem 1.4 from the Lyapunov function introduced. Two proofs are given, one in Sections 3.2 and 3.3 and the other in Section 3.4. The proofs of Theorems 1.3 and 1.5 follow quickly once Theorem 1.4 is proved.

3.1. *Existence and uniqueness of invariant measure.*

PROOF OF PROPOSITION 1.1. Let $\pi$ be the probability measure with density $\rho(x^1,\dots,x^m,y)=\frac{1}{Z}y^{-1/2}$. To prove $\pi=\pi P^\xi$ for $\xi\ll 1$, it suffices to fix an arbitrary state $\bar{\mathbf{x}}=(\bar{x}^1,\dots,\bar{x}^m,\bar{y})\in\mathbf{\Delta}$, let

$$D=D(\bar{\mathbf{x}},\epsilon)=\{\mathbf{x}\in\mathbf{\Delta}\,|\,|x^i-\bar{x}^i|,|y-\bar{y}|<\epsilon\ \forall i\}$$

for $\epsilon>0$ arbitrarily small, and show that

$$\mathbb{P}_\pi[\mathbf{x}_0\in D,E]=\mathbb{P}_\pi[\mathbf{x}_\xi\in D,E]+O(\xi^2),$$

where $E$ is the event that exactly one interaction takes place on the interval $(0,\xi)$. Clearly,

$$\mathbb{P}_\pi[\mathbf{x}_0\in D,E]=\xi(2\epsilon)^m\left(\sum_{i=1}^m\sqrt{\bar{x}^i}\right)\cdot Z^{-1}\bar{y}^{-1/2}+O(\xi^2\epsilon^m+\xi\epsilon^{m+1}).$$

The estimation of $\mathbb{P}_\pi[\mathbf{x}_\xi\in D,E]$ requires a straightforward computation identical to that in Lemma 6.6 of [30]. □

To prove uniqueness, we prove Doeblin's condition on a subset of $\mathbf{\Delta}$, which for convenience we take to be a set of "active states" of the form

$$A_\epsilon:=\{\mathbf{x}\in\mathbf{\Delta}\,|\,x^i,y\geq\epsilon\}$$

for some $\epsilon>0$. For $S\subset\mathbf{\Delta}$, let $U_S$ denote the uniform probability measure on $S$.

PROPOSITION 3.1. *For any* $t>0$ *and* $\epsilon>0$, *there exists a constant* $\eta=\eta(\epsilon,t)$ *such that for every* $\mathbf{x}\in A=A_\epsilon$,

$$P^t(\mathbf{x},\cdot)\geq\eta U_A(\cdot).$$

PROOF. We cover $A$ with finitely many sets of the form $D=D(\bar{\mathbf{x}},\xi)$ where $D(\bar{\mathbf{x}},\xi)$ is as defined in the proof of Proposition 1.1 with the property that $\mathrm{dist}(D,\partial\mathbf{\Delta})>\frac{4\epsilon}{5}$. It suffices to show that given any $t>0$, there exists $\eta>0$ such that for every $\mathbf{x}\in A$, $P^t(\mathbf{x},\cdot)\geq\eta U_D(\cdot)$ for all the $D$ in this cover. There are many ways to arrive at this outcome; below we describe one possible scenario.

Let $\mathbf{x}$ and $D$ be fixed. There will be two rounds of interactions. The first round, which takes place on the time interval $(0,\frac{t}{2})$, will result in most of the energy collecting in the tank; and in the second round, which takes place on $(\frac{t}{2}, t)$, energy is redistributed according to $D$. In more detail, starting from $\mathbf{x}$, the first round consists of particle 1 interacting twice with the tank in quick succession, followed by particle 2, and so on through particle $m$, with no other interactions besides these. For each $i$, the goal of the second interaction is to result in $x^i_{\frac{t}{2}} \in (\frac{2\epsilon}{5}, \frac{3\epsilon}{5})$. This requires two interactions to achieve because after the first interaction, $x^i_{s+} \geq y_s$ [see (1.2)], and tank energy prior to interaction with each particle is $\geq \epsilon$. In the second round, each particle interacts twice with the tank as before, resulting in $x^i_t \in [\bar{x}^i - \xi, \bar{x}^i + \xi]$ uniformly distributed and independent of $x^j_t$ for $j = 1, 2, \ldots, i-1$.

We leave it to the reader to check that the scenario above occurs with probability $\eta > 0$ independent of $\mathbf{x}$ provided $\mathbf{x} \in A$. □

PROOF OF THEOREM 1.2. Let $A = A_\epsilon$ and $t$ be as above. It is obvious that for any $\mathbf{x} \in \mathbf{\Delta}$, $P^{t/2}(\mathbf{x}, A) > 0$. Together with Proposition 3.1, this implies that $P^t(\mathbf{x}, \cdot)$ has a strictly positive density on all of $A$, and that in turn implies that all $\mathbf{x} \in \mathbf{\Delta}$ belong in the same ergodic component, equivalently, $\mathbf{x}_t$ admits at most one invariant probability measure, which must therefore be $\pi$. □

3.2. *Review of tools from probability.* We recall here some tools that we will use to prove polynomial convergence. As these are very general ideas, we will present them in the context of general Markov chains. Let $\Psi_n$ be a (discrete-time) Markov chain on a measurable space $(X, \mathcal{B})$ with transition kernels $\mathcal{P}(x, \cdot)$.

(A) *Atoms of Markov chains.* A set $\alpha \in \mathcal{B}$ is called an *atom* if there is a probability measure $\theta$ on $(X, \mathcal{B})$ such that for all $x \in \alpha$, $\mathcal{P}(x, \cdot) = \theta(\cdot)$. Most Markov chains on continuous or uncountable spaces do not possess atoms. We review here a technique introduced in [35] which shows that under quite general conditions for $\Psi_n$, one can construct explicitly another chain, $\tilde{\Psi}_n$, defined on an enlarged state space $(\tilde{X}, \tilde{\mathcal{B}})$, such that $\tilde{\Psi}_n$ is an extension of $\Psi_n$ and it has an atom.

The relevant condition for $\Psi_n$ is that for some set $A_0 \in \mathcal{B}$, there exists a probability measure $\theta$ and a number $\eta > 0$ such that for every $x \in A_0$, $\mathcal{P}(x, \cdot) \geq \eta\theta(\cdot)$. Let us call a set $A_0$ with this property a *special reference set*. Assuming the existence of such an $A_0$, the splitting technique of [35] is as follows: Let $\tilde{X} = X \cup A_1$ (disjoint union) where $A_1$ is an identical copy of $A_0$, with the obvious extension $\tilde{\mathcal{B}}$ of $\mathcal{B}$ to $\tilde{X}$. First, we define the "lift" of a measure $\mu$ on $(X, \mathcal{B})$ to a measure $\mu^*$ on $(\tilde{X}, \tilde{\mathcal{B}})$:

$$\begin{cases} \mu^*|_X = (1-\eta)\mu|_{A_0} + \mu|_{X \setminus A_0}, \\ \mu^*|_{A_1} = \eta\mu|_{A_0}, \qquad A_0 \cong A_1 \text{ via the natural identification.} \end{cases}$$

The transition kernels $\tilde{\mathcal{P}}(x, \cdot)$ are then given by

$$\begin{cases} \tilde{\mathcal{P}}(x, \cdot) = \big(\mathcal{P}(x, \cdot)\big)^*, & x \in X \setminus A_0, \\ \tilde{\mathcal{P}}(\mathbf{x}, \cdot) = \big[\big(\mathcal{P}(x, \cdot)\big)^* - \eta\theta^*(\cdot)\big]/(1-\eta), & x \in A_0, \\ \tilde{\mathcal{P}}(x, \cdot) = \theta^*(\cdot), & x \in A_1. \end{cases}$$

It is straightforward to check that the chain $\tilde{\Psi}_n$ projects to $\Psi_n$, meaning $(\mu\mathcal{P})^* = \mu^*\tilde{\mathcal{P}}$, so that $\|\mu\mathcal{P}^n - \nu\mathcal{P}^n\|_{\mathrm{TV}} \leq \|\mu^*\tilde{\mathcal{P}}^n - \nu^*\tilde{\mathcal{P}}^n\|_{\mathrm{TV}}$. Finally, $A_1$ is an atom for the chain $\tilde{\Psi}_n$—this is the whole point of the construction.

(B) *Connection to renewal processes*. For $E \in \mathcal{B}$, we let $\tau_E$ denote the first passage time to $E$, that is,

$$\tau_E = \inf\{n > 0 | \Psi_n \in E\}.$$

Suppose the chain $\Psi_n$ has an atom $\alpha$, and that $\alpha$ is accessible, that is, $\mathcal{P}_x[\tau_\alpha < \infty] = 1$ for every $x \in X$. Given two initial distributions $\mu$ and $\nu$ on $X$, we wish to bound the rate at which $\|\mu\mathcal{P}^n - \lambda\mathcal{P}^n\|_{\mathrm{TV}}$ tends to 0 as $n \to \infty$ where $\|\cdot\|_{\mathrm{TV}}$ is the total variational norm. One way to proceed is to run two independent copies of the chain with initial distributions $\mu$ and $\nu$ respectively, and perform a coupling at simultaneous returns to the atom $\alpha$. It is well known that if $T$ is the coupling time, then

$$\|\mu\mathcal{P}^n - \nu\mathcal{P}^n\|_{\mathrm{TV}} \leq 2\mathbb{P}[T > n]. \tag{3.1}$$

The quantities $\mathbb{P}[T > n]$, on the other hand, can be studied via two associated renewal processes as follows.

Let $Y_0$ and $Y_0'$ be independent $\mathbb{N}$-valued random variables having the distributions of $\tau_\alpha$, the first passage time to $\alpha$, starting from $\mu$ and $\nu$, respectively, and let $Y_1, Y_2, \dots$ and $Y_1', Y_2', \dots$ be *i.i.d.* random variables the distributions of which are equal to that of $\tau_\alpha$ starting from $\alpha$. In addition, we assume the *return times* to $\alpha$ are *aperiodic*, that is, $\gcd\{n \geq 1 | \mathbb{P}[Y_i = n] > 0\} = 1$. Then $S_n := \sum_{i=0}^n Y_i$ and $S_n' := \sum_{i=0}^n Y_i'$, $n = 0, 1, 2, \dots$, are renewal processes, and $T$ above is the first simultaneous renewal time, that is,

$$T = \inf_{n \geq 0}\{S_{k_1} = S'_{k_2} = n \text{ for some } k_1, k_2\}.$$

The following known result relates the finiteness of the moments of $T$ to the corresponding moments for the distributions of $Y_0$, $Y_0'$ and $Y_1$.

THEOREM 3.2 (Theorem 4.2 of [32]). *Let $Y_i$ and $Y_i'$ be as above. Suppose that for some $\beta \geq 1$, we have*

$$\mathbb{E}[Y_0^\beta], \qquad \mathbb{E}[Y_0'^\beta] \quad \textit{and} \quad \mathbb{E}[Y_1^\beta] < \infty. \tag{3.2}$$

*Then $\mathbb{E}[T^\beta]$ is also finite.*

The discussion above implies the following.

COROLLARY 3.3. *Let $\Psi_n$ be a Markov chain on $(X, \mathcal{B})$ with transition kernel $\mathcal{P}$. Suppose $\Psi_n$ has an atom $\alpha$ that is accessible and whose return times are* *aperiodic. Let $\mu$ and $\nu$ be two probability distributions on $X$, and assume that for some $\beta > 1$,*

$$\mathbb{E}_\mu[\tau_\alpha^\beta], \qquad \mathbb{E}_\nu[\tau_\alpha^\beta] \quad \textit{and} \quad \mathbb{E}_\alpha[\tau_\alpha^\beta] < \infty.$$

*Then*

$$\lim_{n\to\infty} n^\beta \|\mu\mathcal{P}^n - \nu\mathcal{P}^n\|_{\mathrm{TV}} = 0.$$

The proof is as discussed, together with the following general relation: Let $Z$ be a random variable taking values in $\mathbb{N}$, and let $\beta > 1$. Then

$$\mathbb{E}[Z^\beta] < \infty \quad \Longrightarrow \quad \lim_{n\to\infty} n^\beta \mathbb{P}[Z > n] = 0. \tag{3.3}$$

(C) *Lyapunov function and moments of first passage times.* The following result, which is sufficient for our purposes, is a simple version of Theorem 3.6 of [22].

THEOREM 3.4 (Theorem 3.6 of [22]). *Let $\Psi_n$ be a Markov chain on $(X, \mathcal{B})$ with transition kernel $\mathcal{P}$. We assume that there exist a function $W : X \to [1, \infty)$, a set $A \in \mathcal{B}$, constants $b, c > 0$ and $0 \le \beta < 1$ such that*

$$\mathcal{P}W - W \le -cW^\beta + b\mathbf{1}_A. \tag{3.4}$$

*Then there is a constant $\hat{c}$ such that for all $x \in X$,*

$$\mathbb{E}_x\left[\sum_{k=0}^{\tau_A - 1} (k+1)^{\hat{\beta}-1}\right] \le \hat{c}W(x), \qquad \hat{\beta} = (1-\beta)^{-1}.$$

Clearly, $\mathbb{E}_x[\tau_A^{\hat{\beta}}]$ is bounded above by a constant times the expectation above.

The reader may notice that we have omitted some of the hypotheses in Theorem 3.6 of [22] in the statement of Theorem 3.4 above. This is because they are not needed: here we consider only the first passage time to $A$, which can be thought of as a set of the form $\{W \le \text{constant}\}$, while [22] considers first passage times to arbitrary sets. We remark also that [22] does not give the rate of convergence to equilibrium we claim; it shows that in general, convergence rate is bounded by $\sim t^{\hat{\beta}-1}$, but as we will see, additional information for our systems enables us to prove a faster convergence rate $\sim t^{\hat{\beta}-2}$.

REMARKS. In (A), (B) and (C) above, we have outlined a general strategy for deducing polynomial rates of convergence or of correlation decay for Markov chains. While we have cited specific references, they are not the only ones that contributed to this general body of ideas [8, 9, 14, 20, 36, 41]. We acknowledge in particular [36], which was proved earlier and which used similar ideas as above though some of the arguments were carried out a little differently. We mention also [48], which models deterministic dynamical systems with chaotic behavior as objects that are slight generalizations of countable state Markov chains. This paper focuses on tails of return times, that is, $\mathbb{P}[\tau_\alpha > n]$, rather than on moments of $\tau_\alpha$, to a set $\alpha$ that is effectively a special reference set as defined in (A); tails of first passage times and moments are, as we have noted, essentially equivalent.

3.3. *Proofs of theorems.* We first prove Theorem 1.4. Theorems 1.3 and 1.5 follow easily; their proofs are given at the end of the subsection.

Let $h > 0$ be small enough for Theorem 2.1 to apply, and let

$$\hat{\mathbf{x}}_n = (\hat{x}_n^1, \dots, \hat{x}_n^m, \hat{y}_n) = (x_{nh}^1, \dots, x_{nh}^m, y_{nh}), \qquad n = 0, 1, 2, \dots,$$

be the time-$h$ sampling chain of $\mathbf{x}_t$. Letting $\lfloor \frac{t}{h} \rfloor$ denote the largest integer $\leq \frac{t}{h}$, we observe that

$$\begin{aligned} \|\mu P^t - \nu P^t\|_{\mathrm{TV}} &= \|(\mu P^{\lfloor \frac{t}{h} \rfloor h} - \nu P^{\lfloor \frac{t}{h} \rfloor h}) P^{(t - \lfloor \frac{t}{h} \rfloor)h}\|_{\mathrm{TV}} \\ &\leq \|\mu P^{\lfloor \frac{t}{h} \rfloor h} - \nu P^{\lfloor \frac{t}{h} \rfloor h}\|_{\mathrm{TV}}, \end{aligned}$$

so it suffices to prove the theorem for $\hat{\mathbf{x}}_n$ corresponding to a fixed $h$. From here on, $h$ is fixed, and since we will be working exclusively with the discrete-time chain $\hat{\mathbf{x}}_n$, the $\hat{}$ in $\hat{\mathbf{x}}_n$ is dropped for notational simplicity.

Let $\gamma > 0$ be small enough that Theorem 2.1 applies with $\alpha = \frac{1}{2} - \frac{\gamma}{8}$. We define

$$\mathcal{A} = \mathcal{A}_{\gamma,h} = \{\mathbf{x} \in \mathbf{\Delta} | V_{\frac{1}{2} - \frac{\gamma}{8}}(\mathbf{x}) \leq M\},$$

where $M = M(\frac{1}{2} - \frac{\gamma}{8}, h)$, and let

$$\tau_{\mathcal{A}} = \inf_{n > 0}\{\mathbf{x}_n \in \mathcal{A}\}$$

be the first passage time to $\mathcal{A}$. We plan to proceed as follows:

(1) First we estimate the moments of $\tau_{\mathcal{A}}$.

(2) Using $\mathcal{A}$ as a special reference set, we split the chain, obtaining an atom $\alpha$ for the split chain $\tilde{\mathbf{x}}_n$.

(3) We deduce from (1) the moments of $\tilde{\tau}_\alpha$, the first passage time of $\tilde{\mathbf{x}}_n$ to $\alpha$, and

(4) finally, we apply Corollary 3.3 to $\tilde{\tau}_\alpha$ to obtain the desired results.

LEMMA 3.5. *Given $\gamma$ as above, there exists $C = C(\gamma)$ such that for all $\mathbf{x} \in \mathbf{\Delta}$,*

$$\mathbb{E}_{\mathbf{x}}[\tau_{\mathcal{A}}^{2 - \frac{\gamma}{2}}] \leq C V_{\frac{1}{2} - \frac{\gamma}{8}}(\mathbf{x}).$$

PROOF. We apply Theorem 2.1 to $V_{\frac{1}{2} - \frac{\gamma}{8}}$. From Corollary 2.3, it follows that if $W = \max\{V_{\frac{1}{2} - \frac{\gamma}{8}}, 1\}$, then $b := \sup_{\mathbf{x} \in \mathcal{A}}\{P^h W(\mathbf{x}) - W(\mathbf{x})\} < \infty$, and we have

$$P^h W - W \leq -c W^{1 - \frac{2}{4 - \gamma}} + b \mathbf{1}_{\mathcal{A}}.$$

Theorem 3.4 then tells us that there is a constant $\hat{c}$ such that for all $\mathbf{x}$,

$$\mathbb{E}_{\mathbf{x}}\left[\sum_{k=0}^{\tau_{\mathcal{A}}-1}(k+1)^{1-\frac{\gamma}{2}}\right] \leq \hat{c}W(\mathbf{x}).$$

As $W \leq C_2 V_{\frac{1}{2}-\frac{\gamma}{8}}$ for some constant $C_2 > 0$ that depends only on $\bar{E}, m$ and $\gamma$, it follows that

$$\mathbb{E}_{\mathbf{x}}[\tau_{\mathcal{A}}^{2-\frac{\gamma}{2}}] \leq 2 \cdot \mathbb{E}_{\mathbf{x}}\left[\sum_{k=0}^{\tau_{\mathcal{A}}-1}(k+1)^{1-\frac{\gamma}{2}}\right] \leq 2C_2 \cdot \hat{c} \cdot V_{\frac{1}{2}-\frac{\gamma}{8}}(\mathbf{x}).$$

This completes the proof. □

Recall that for small $\delta > 0$, $\mathcal{M}_\delta$ is the set of Borel probability measures $\mu$ on $\mathbf{\Delta}$ such that

$$\int_{\mathbf{\Delta}}\left(\sum_{k=1}^{m}(x^k)^{2\delta-1} + y^{\delta-\frac{1}{2}}\right)\mu(\mathrm{d}\mathbf{x}) \equiv \int_{\mathbf{\Delta}} V_{\frac{1}{2}-\delta}(\mathbf{x})\mu(\mathrm{d}\mathbf{x}) < \infty.$$

PROOF OF THEOREM 1.4. Let $\gamma > 0$ and $h > 0$ be as above, and let $\mu, \nu \in \mathcal{M}_{\gamma/8}$ be given. It follows from Lemma 3.5 that

$$\mathbb{E}_\mu[\tau_{\mathcal{A}}^{2-\frac{\gamma}{2}}], \mathbb{E}_\nu[\tau_{\mathcal{A}}^{2-\frac{\gamma}{2}}] < \infty.$$

Observe next that $\mathcal{A}$ is a special reference set in the sense of Section 3.2(A); this follows from Proposition 3.1, for $\mathcal{A} \subset A_\epsilon$ for $\epsilon > 0$ small enough. We split the chain as discussed in Section 3.2(A), denoting the split chain by $\tilde{\mathbf{x}}_n$, and let $\mathcal{A}_0$ and $\mathcal{A}_1$ be identical copies of $\mathcal{A}$ in $\tilde{\mathbf{\Delta}}$, with $\mathcal{A}_1 = \alpha$ being an atom.

To apply Corollary 3.3 to the chain $\tilde{\mathbf{x}}_n$, we first check that the atom $\alpha$ is accessible: It is easy to see that if $\tilde{\tau}$ is first passage time of $\tilde{\mathbf{x}}_n$, then $\tau_{\mathcal{A}} = \tilde{\tau}_{\mathcal{A}_0 \cup \mathcal{A}_1}$, and from Theorem 1.2, we know that $\mathcal{A}$ is accessible under $\mathbf{x}_n$. Moreover, every time $\tilde{\mathbf{x}}_n$ returns to $\mathcal{A}_0 \cup \mathcal{A}_1$, it has probability $\eta$ of entering $\alpha$. This guarantees the accessibility of $\alpha$. Aperiodicity of return times to $\alpha$ follows from the fact that for all $\tilde{\mathbf{x}}_0$, $\mathbb{P}[\tilde{\mathbf{x}}_1 \in \alpha] > 0$.

It remains to pass the moments of $\tau_{\mathcal{A}}$ to the moments $\tilde{\tau}_\alpha$. For a measure $\lambda$ on $\mathbf{\Delta}$, $\tilde{\lambda}$ denotes its lift to $\tilde{\mathbf{\Delta}}$.

LEMMA 3.6. (i) $\mathbb{E}_\alpha[\tilde{\tau}_\alpha^{2-\gamma}] < \infty$.
(ii) $\mathbb{E}_{\tilde{\lambda}}[\tilde{\tau}_\alpha^{2-\gamma}] < \infty$ *for* $\lambda$ *with* $\mathbb{E}_\lambda[\tau_{\mathcal{A}}^{2-\frac{\gamma}{2}}] < \infty$.

This lemma follows from Lemma 3.1 of [36]; we provide an elementary proof below for completeness. Assuming Lemma 3.6, we may now apply Corollary 3.3 to $\tilde{\mathbf{x}}_n$ with $\beta = 2 - \gamma$, giving a convergence rate of $t^{2-\gamma}$. To finish, recall from Section 3.2(A) that if $\mathcal{P}$ and $\tilde{\mathcal{P}}$ are the Markov operators for $\mathbf{x}_n$ and $\tilde{\mathbf{x}}_n$, respectively, then $\|\mu\mathcal{P}^n - \nu\mathcal{P}^n\|_{\mathrm{TV}} \leq \|\tilde{\mu}\tilde{\mathcal{P}}^n - \tilde{\nu}\tilde{\mathcal{P}}^n\|_{\mathrm{TV}}$. □

PROOF OF LEMMA 3.6. To prove (i), it suffices to show that for some $\gamma' < \gamma$, there exists $C$ such that

$$\mathbb{P}_\alpha[\tilde{\tau}_\alpha > k] \le Ck^{\gamma'-2} \qquad \text{for all } k. \tag{3.5}$$

Let $\tau_n, n = 1, 2, \dots$, denote the $n$th entrance time into $\mathcal{A}_0 \cup \mathcal{A}_1$, and let $\mathbf{N}$ be smallest $n$ such that $\tilde{\tau}_\alpha = \tau_n$. Since at each $\tau_n$, the probability of being in $\alpha$ is $\eta$, we have $\mathbb{P}[\mathbf{N} > k] = (1-\eta)^k$. Note also that since $\sup_{\mathbf{x}\in\mathcal{A}} \mathbb{E}_\mathbf{x}[\tau_\mathcal{A}^{2-\frac{\gamma}{2}}] < \infty$ by Lemma 3.5, it follows that

$$\mathbb{P}[\tau_{n+1} - \tau_n \ge k | \mathbf{N} > n, \tilde{\mathbf{x}}_{\tau_n}] \le C' k^{-(2-\frac{\gamma}{2})}$$

for some constant $C'$.

For any $\delta > 0$, we have

$$\{\tau_\mathbf{N} > k^{1+\delta}\} \subset \{\mathbf{N} > k^\delta\} \cup \bigcup_{n=0}^{\lfloor k^\delta \rfloor} \{\tau_{n+1} - \tau_n > k; \mathbf{N} > n\}.$$

Thus,

$$\begin{aligned}
\mathbb{P}_\alpha[\tilde{\tau}_\alpha > k^{1+\delta}] &\le \mathbb{P}_\alpha[\mathbf{N} > k^\delta] + \sum_{n=0}^{\lfloor k^\delta \rfloor} \mathbb{P}_\alpha[\tau_{n+1} - \tau_n > k | \mathbf{N} > n] \\
&= \mathbb{P}_\alpha[\mathbf{N} > k^\delta] \\
&\quad + \sum_{n=0}^{\lfloor k^\delta \rfloor} \int \mathbb{P}_\alpha[\tau_{n+1} - \tau_n > k | \mathbf{N} > n, \tilde{\mathbf{x}}_{\tau_n} = \tilde{\mathbf{x}}] \mathbb{P}_\alpha[\tilde{\mathbf{x}}_{\tau_n} = d\tilde{\mathbf{x}}, \mathbf{N} > n] \\
&\le (1-\eta)^{k^\delta} + k^\delta C' k^{-(2-\frac{\gamma}{2})}.
\end{aligned}$$

Noting that the second term dominates for large $k$, we obtain (3.5) by choosing $\delta$ sufficiently small.

The proof of (ii) follows similar steps and uses the finiteness of $\mathbb{E}_\mu[\tau_\mathcal{A}^{2-\frac{\gamma}{2}}]$. $\square$

PROOF OF THEOREM 1.3. A simple computation using the density of $\pi$ shows that $\pi \in \mathcal{M}_\delta$ for every $\delta > 0$. Also, for every $\mathbf{x} \in \mathbf{\Delta}$, the point mass $\delta_\mathbf{x}$ clearly belongs in $\mathcal{M}_\delta$ for all $\delta > 0$. Thus, Theorem 1.3 is a special case of Theorem 1.4, with $\mu = \pi$ and $\nu = \delta_\mathbf{x}$. $\square$

PROOF OF THEOREM 1.5. As a direct consequence of Theorem 1.4, we have

$$\begin{aligned}
&\left| \int (P^t\zeta)(\mathbf{x})\xi(\mathbf{x})\mu(d\mathbf{x}) - \int (P^t\zeta)(\mathbf{x})\mu(d\mathbf{x}) \int \xi(\mathbf{x})\mu(d\mathbf{x}) \right| \\
&\quad = \left| \int \xi(\mathbf{x}) \Big( (P^t\zeta)(\mathbf{x}) - \int (P^t\zeta)(\mathbf{z})\mu(d\mathbf{z}) \Big) \mu(d\mathbf{x}) \right|
\end{aligned}$$

$$\leq \|\xi\|_{L^\infty}\|\zeta\|_{L^\infty}\int \|\delta_{\mathbf{x}}P^t - \mu P^t\|_{\mathrm{TV}}\mu(\mathrm{d}\mathbf{x})$$

$$= o\left(\frac{1}{t^{2-\gamma}}\right). \qquad \square$$

PROOF OF THEOREM 1.6. Theorem 1.6 follows straightforwardly from the following Markov chain central limit theorem (Corollary 2 of [23]):

*Suppose $\Psi_n$ is a Harris ergodic Markov chain on $(X, \mathcal{B})$ with transition kernel $\mathcal{P}$. Let $\mu$ be the stationary distribution of $\Psi_n$ and let $f : X \to \mathbb{R}$ be a Borel function that is uniformly bounded $\mu$-almost surely. Assume $\Psi_n$ is polynomial ergodic such that*

$$\|\mathcal{P}^n(x,\cdot) - \mu\|_{\mathrm{TV}} \leq M(x)n^{-m},$$

*where $m > 1$ and $\mathbb{E}_\pi(M) < \infty$, then for any initial distribution, as $n \to \infty$,*

$$\sqrt{n}(\bar{f} - \mathbb{E}_\mu f) \xrightarrow{d} N(0, \sigma_f^2),$$

*where*

$$\sigma_f^2 := \mathrm{var}\{f(\Psi_0)\} + 2\sum_{i=1}^{\infty}\mathrm{cov}\{f(\Psi_0), f(\Psi_i)\} < \infty.$$

It is a simple exercise to check that all conditions are satisfied by the time-$\delta$ chain $\{\mathbf{x}_{n\delta}\}_{n=0}^{\infty}$ for any $\delta > 0$. $\square$

3.4. *Alternate proof of Theorem* 1.4. As pointed out by one of our reviewers, Theorem 1.4 also follows from Theorem 4.1 in [20]. We thank the reviewer for pointing us to this result. Below we recall the statement of it, and then show how to use it to deduce Theorem 1.4.

Let $\Psi_t$ be a strong Markov chain on a metric space $X$ with infinitesimal generator $\mathcal{L}$ and associated semigroup $\mathcal{P}_t$. The following result of sub-geometric rates of convergence holds.

THEOREM 3.7 (Theorem 4.1 of [20]). *Assume $\Psi_t$ has a cadlag modification and $\mathcal{P}_t$ is Feller. Assume furthermore that there exists a continuous function $V : X \mapsto [1,\infty)$ with pre-compact sublevel sets such that*

$$\mathcal{L}V \leq K - \phi(V)$$

*for some constant $K$ and for some strictly concave function $\phi : \mathbb{R}^+ \to \mathbb{R}^+$ with $\phi(0) = 0$ and increasing to infinity. In addition, we assume that sublevel sets of $V$ are "small" in the sense that for every $C > 0$ there exists $\alpha > 0$ and $T > 0$ such that*

$$\|\mathcal{P}_T(x,\cdot) - \mathcal{P}_T(y,\cdot)\|_{\mathrm{TV}} \leq 2(1-\alpha)$$

*for every $(x, y)$ such that $V(x) + V(y) \leq C$. Then*:

- *There exists a unique invariant measure $\mu$ for $\Psi_t$ and $\mu$ is such that*

$$\int_X \phi(V(x))\mu(\mathrm{d}x) \leq K.$$

- *Let $H_\phi$ be the function defined by*

$$H_\phi(u) = \int_1^u \frac{\mathrm{d}s}{\phi(s)}.$$

*Then there exists a constant $C$ such that for every $x, y \in X$, one has the bounds*

$$\|\mathcal{P}_t(x,\cdot) - \mathcal{P}_t(y,\cdot)\|_{\mathrm{TV}} \leq C \frac{V(x)+V(y)}{H_\phi^{-1}(t)}.$$

The proof of Theorem 4.1 uses a different coupling that bypasses the explicit splitting of the Markov chain, and the Lyapunov function is lifted to $X \times X$. Similar estimates of hitting times as in Lemmas 3.5 and 3.6 are also ingredients in this proof.

PROOF OF THEOREM 1.4 USING THEOREM 4.1. It is a simple exercise to check that (1) $\mathbf{x}_t$ is a strong Markov process with an infinitesimal generator $\mathcal{G}$, and (2) $\mathbf{x}_t$ is a Feller process with cadlag sample paths.

Let $V(\mathbf{x}) = V_\alpha(\mathbf{x})$ be the same Lyapunov function used before. (One may multiply $V$ by a constant to make its minimum be greater than 1, if necessary.) We have

$$\mathcal{G}V(\mathbf{x}) = \sum_{i=1}^m Q_i,$$

where

$$Q_i = \sqrt{x_0^i}\left\{\int_0^1 [(x_0^i(1-u^2)+y_0)^{-2\alpha} + (x_0^i u^2)^{-\alpha}]\,\mathrm{d}u - [(x_0^i)^{-2\alpha} + y_0^{-\alpha}]\right\}.$$

Therefore it follows from Lemma 2.2 that there exist constants $\epsilon_0 > 0$ and $c^* > 0$ such that

$$\mathcal{G}V(\mathbf{x}) \leq -c^* V^\beta(\mathbf{x})$$

for every $\mathbf{x} \in B$, where $\beta = 1 - (4\alpha)^{-1}$ and

$$B = \{\mathbf{x} \in \mathbf{\Delta} | y < \epsilon_0,\ \text{or } x^i < 4^{-\frac{1}{2\alpha}}\epsilon_0 \text{ for some } i \in \{1,\dots,m\}\}.$$

Let

$$K = \sup_{\mathbf{x}\in\mathbf{\Delta}\setminus B} \sum_{i=1}^m Q_i(\mathbf{x}).$$

It is easy to check that $K < \infty$ and

$$\mathcal{G}V \le K - c^* V^\beta.$$

It remains to check that the sublevel sets of $V$ are "small." Let $A = A_C$ be the sublevel set $\{V \le C\}$. By the same proof as in Proposition 3.1, for any $t > 0$ and $C > \min V(\mathbf{x})$, there exists a constant $\eta = \eta(C, t)$ such that for every $\mathbf{x} \in A$,

$$P^t(\mathbf{x}, \cdot) \ge \eta U_A(\cdot),$$

where $U_A$ is the uniform probability measure on $A$. This implies

$$\|P^t(\mathbf{x}, \cdot) - P^t(\mathbf{y}, \cdot)\|_{\mathrm{TV}} \le 2(1 - \eta)$$

for each $\mathbf{x}, \mathbf{y}$ such that $V(\mathbf{x}) + V(\mathbf{y}) \le C$.

Therefore, let $\phi(x) = c^* x^\beta$, by Theorem 4.1, we have

$$\|P^t(\mathbf{x}, \cdot) - P^t(\mathbf{y}, \cdot)\|_{\mathrm{TV}} \le C_0 \frac{V(\mathbf{x}) + V(\mathbf{y})}{(c^* \int_1^t s^{(4\alpha)^{-1}-1}\, \mathrm{d}s)^{-1}} = C_0' \big(V(\mathbf{x}) + V(\mathbf{y})\big) t^{-4\alpha}$$

for some constants $C_0$ and $C_0'$. The proof of Theorem 1.4 is completed by letting $\alpha = \frac{1}{2} - \gamma/4$. □

## REFERENCES


[1] BALAJI, S. and MEYN, S. P. (2000). Multiplicative ergodicity and large deviations for an irreducible Markov chain. *Stochastic Process. Appl.* **90** 123–144. MR1787128

[2] BOWEN, R. (2008). *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms*, revised ed. *Lecture Notes in Math.* **470**. Springer, Berlin. MR2423393

[3] BUNIMOVICH, L., LIVERANI, C., PELLEGRINOTTI, A. and SUHOV, Y. (1992). Ergodic systems of *n* balls in a billiard table. *Comm. Math. Phys.* **146** 357–396. MR1165188

[4] CARLEN, E. A., CARVALHO, M. C. and LOSS, M. (2003). Determination of the spectral gap for Kac's master equation and related stochastic evolution. *Acta Math.* **191** 1–54. MR2020418

[5] CARLEN, E. A., CARVALHO, M. C. and LOSS, M. (2014). Spectral gap for the Kac model with hard sphere collisions. *J. Funct. Anal.* **266** 1787–1832. MR3146836

[6] CHERNOV, N. and MARKARIAN, R. (2006). *Chaotic Billiards*. *Mathematical Surveys and Monographs* **127**. Amer. Math. Soc., Providence, RI. MR2229799

[7] CHERNOV, N. and YOUNG, L. S. (2000). Decay of correlations for Lorentz gases and hard balls. In *Hard Ball Systems and the Lorentz Gas*. *Encyclopaedia Math. Sci.* **101** 89–120. Springer, Berlin. MR1805327

[8] DOUC, R., FORT, G. and GUILLIN, A. (2009). Subgeometric rates of convergence of *f*-ergodic strong Markov processes. *Stochastic Process. Appl.* **119** 897–923. MR2499863

[9] DOUC, R., FORT, G., MOULINES, E. and SOULIER, P. (2004). Practical drift conditions for subgeometric rates of convergence. *Ann. Appl. Probab.* **14** 1353–1377. MR2071426

[10] ECKMANN, J.-P. and JACQUET, P. (2007). Controllability for chains of dynamical scatterers. *Nonlinearity* **20** 1601–1617. MR2335075

[11] ECKMANN, J.-P. and MEJÍA-MONASTERIO, C. (2006). Thermal rectification in billiardlike systems. *Phys. Rev. Lett.* **97** 094301.

[12] ECKMANN, J.-P., MEJÍA-MONASTERIO, C. and ZABEY, E. (2006). Memory effects in nonequilibrium transport for deterministic Hamiltonian systems. *J. Stat. Phys.* **123** 1339–1360. MR2253882

[13] ECKMANN, J.-P. and YOUNG, L.-S. (2006). Nonequilibrium energy profiles for a class of 1-D models. *Comm. Math. Phys.* **262** 237–267. MR2200889

[14] FORT, G. and ROBERTS, G. O. (2005). Subgeometric ergodicity of strong Markov processes. *Ann. Appl. Probab.* **15** 1565–1589. MR2134115

[15] GASPARD, P. and GILBERT, T. (2008). Heat conduction and Fourier's law in a class of many particle dispersing billiards. *New J. Phys.* **10** 103004.

[16] GASPARD, P. and GILBERT, T. (2008). Heat conduction and Fourier's law by consecutive local mixing and thermalization. *Phys. Rev. Lett.* **101** 020601.

[17] GASPARD, P. and GILBERT, T. (2008). On the derivation of Fourier's law in stochastic energy exchange systems. *J. Stat. Mech. Theory Exp.* **2008** P11021.

[18] GAVEAU, B. and KAC, M. (1986). A probabilistic formula for the quantum $N$-body problem and the nonlinear Schrödinger equation in operator algebra. *J. Funct. Anal.* **66** 308–322. MR0839104

[19] GRIGO, A., KHANIN, K. and SZÁSZ, D. (2012). Mixing rates of particle systems with energy exchange. *Nonlinearity* **25** 2349–2376. MR2956578

[20] HAIRER, M. (2010). Convergence of Markov processes. *Lecture Notes*. Available at http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.728.1839.

[21] JANVRESSE, E. (2001). Spectral gap for Kac's model of Boltzmann equation. *Ann. Probab.* **29** 288–304. MR1825150

[22] JARNER, S. F. and ROBERTS, G. O. (2002). Polynomial convergence rates of Markov chains. *Ann. Appl. Probab.* **12** 224–247. MR1890063

[23] JONES, G. L. (2004). On the Markov chain central limit theorem. *Probab. Surv.* **1** 299–320. MR2068475

[24] KAC, M. (1956). Foundations of kinetic theory. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, 1954–1955, *Vol. III* 171–197. Univ. California Press, Berkeley. MR0084985

[25] KHANIN, K. and YARMOLA, T. (2013). Ergodic properties of random billiards driven by thermostats. *Comm. Math. Phys.* **320** 121–147. MR3046992

[26] KONTOYIANNIS, I. and MEYN, S. P. (2005). Large deviations asymptotics and the spectral theory of multiplicatively regular Markov processes. *Electron. J. Probab.* **10** 61–123 (electronic). MR2120240

[27] LEFEVERE, R., MARIANI, M. and ZAMBOTTI, L. (2011). Large deviations of the current in stochastic collisional dynamics. *J. Math. Phys.* **52** 033302, 22. MR2814865

[28] LI, Y. (2015). On the stochastic behaviors of locally confined particle systems. *Chaos* **25** 073121, 14. MR3405854

[29] LI, Y. and YOUNG, L.-S. (2013). Existence of nonequilibrium steady state for a simple model of heat conduction. *J. Stat. Phys.* **152** 1170–1193. MR3101481

[30] LI, Y. and YOUNG, L.-S. (2014). Nonequilibrium steady states for a class of particle systems. *Nonlinearity* **27** 607–636. MR3177574

[31] LIN, K. K. and YOUNG, L.-S. (2010). Nonequilibrium steady states for certain Hamiltonian models. *J. Stat. Phys.* **139** 630–657. MR2638931

[32] LINDVALL, T. (2002). *Lectures on the Coupling Method*. Dover Publications, Mineola, NY. MR1924231

[33] MASLEN, D. K. (2003). The eigenvalues of Kac's master equation. *Math. Z.* **243** 291–331. MR1961868

[34] MEJIA-MONASTERIO, C., LARRALDE, H. and LEYVRAZ, F. (2001). Coupled normal heat and matter transport in a simple model system. *Phys. Rev. Lett.* **86** 5417.

[35] NUMMELIN, E. (1978). A splitting technique for Harris recurrent Markov chains. *Z. Wahrsch. Verw. Gebiete* **43** 309–318. MR0501353

[36] NUMMELIN, E. and TUOMINEN, P. (1983). The rate of convergence in Orey's theorem for Harris recurrent Markov chains with applications to renewal theory. *Stochastic Process. Appl.* **15** 295–311. MR0711187

[37] RATEITSCHAK, K., KLAGES, R. and NICOLIS, G. (2000). Thermostating by deterministic scattering: The periodic Lorentz gas. *J. Stat. Phys.* **99** 1339–1364. MR1773148

[38] REY-BELLET, L. and YOUNG, L.-S. (2008). Large deviations in non-uniformly hyperbolic dynamical systems. *Ergodic Theory Dynam. Systems* **28** 587–612. MR2408394

[39] SASADA, M. (2015). Spectral gap for stochastic energy exchange model with nonuniformly positive rate function. *Ann. Probab.* **43** 1663–1711. MR3353812

[40] SINAĬ, JA. G. (1970). Dynamical systems with elastic reflections. Ergodic properties of dispersing billiards. *Uspekhi Mat. Nauk* **25** 141–192. MR0274721

[41] TUOMINEN, P. and TWEEDIE, R. L. (1994). Subgeometric rates of convergence of $f$-ergodic Markov chains. *Adv. in Appl. Probab.* **26** 775–798. MR1285459

[42] WU, L. (2000). Uniformly integrable operators and large deviations for Markov processes. *J. Funct. Anal.* **172** 301–376. MR1753178

[43] WU, L. (2001). Large and moderate deviations and exponential convergence for stochastic damping Hamiltonian systems. *Stochastic Process. Appl.* **91** 205–238. MR1807683

[44] YARMOLA, T. (2011). Ergodicity of some open systems with particle–disk interactions. *Comm. Math. Phys.* **304** 665–688. MR2794543

[45] YARMOLA, T. (2013). Sub-exponential mixing of random billiards driven by thermostats. *Nonlinearity* **26** 1825–1837. MR3071443

[46] YARMOLA, T. (2014). Sub-exponential mixing of open systems with particle–disk interactions. *J. Stat. Phys.* **156** 473–492. MR3217533

[47] YOUNG, L.-S. (1998). Statistical properties of dynamical systems with some hyperbolicity. *Ann. of Math.* (2) **147** 585–650. MR1637655

[48] YOUNG, L.-S. (1999). Recurrence times and rates of mixing. *Israel J. Math.* **110** 153–188. MR1750438

DEPARTMENT OF MATHEMATICS AND STATISTICS
UNIVERSITY OF MASSACHUSETTS AMHERST
AMHERST, MASSACHUSETTS 01003
USA
E-MAIL: yaoli@math.umass.edu

COURANT INSTITUTE OF MATHEMATICAL SCIENCES
NEW YORK UNIVERSITY
NEW YORK, NEW YORK 10012
USA
E-MAIL: lsy@cims.nyu.edu