Ergod. Th. & Dynam. Sys. (2017), **37**, 2556–2596 © *Cambridge University Press, 2016*
doi:10.1017/etds.2016.9

Quasistatic dynamical systems

NEIL DOBBS† and MIKKO STENLUND‡

† *Département de Physique Théorique, Université de Genève, Geneva 1211, Switzerland*
(e-mail: neil.dobbs@gmail.com)
‡ *Department of Mathematics and Statistics, P.O. Box 68,*
Fin-00014 University of Helsinki, Finland
(e-mail: mikko.stenlund@helsinki.fi)

(*Received* 11 *May* 2015 *and accepted in revised form* 11 *January* 2016)

Abstract. We introduce the notion of a quasistatic dynamical system, which generalizes that of an ordinary dynamical system. Quasistatic dynamical systems are inspired by the namesake processes in thermodynamics, which are idealized processes where the observed system transforms (infinitesimally) slowly due to external influence, tracing out a continuous path of thermodynamic equilibria over an (infinitely) long time span. Time evolution of states under a quasistatic dynamical system is entirely deterministic, but choosing the initial state randomly renders the process a stochastic one. In the prototypical setting where the time evolution is specified by strongly chaotic maps on the circle, we obtain a description of the statistical behavior as a stochastic diffusion process, under surprisingly mild conditions on the initial distribution, by solving a well-posed martingale problem. We also consider various admissible ways of centering the process, with the curious conclusion that the 'obvious' centering suggested by the initial distribution sometimes fails to yield the expected diffusion.

1. *Introduction*

This paper belongs to the wider context of studying statistical properties of dynamical systems using probabilistic techniques. We are going to propose a new, but very natural, class of *non-stationary* dynamical systems, and to implement methods from probability theory (e.g. coupling) and stochastic analysis (a martingale problem) in their investigation, alongside methods from ergodic theory and dynamical systems.

1.1. *Motivation and abstract set-up.*

This paper grew out of the will to understand the following abstract set-up. Consider a system $\mathfrak{s}$ in contact with an ambient system $\mathfrak{S}$. The joint system $(\mathfrak{s}, \mathfrak{S})$ is too large and complex to analyze, while $\mathfrak{s}$, the observed subsystem of actual interest, is more amenable to scrutiny. The observed system $\mathfrak{s}$ and the ambient

https://doi.org/10.1017/etds.2016.9 Published online by Cambridge University Press

system $\mathfrak{S}$ are allowed to interact, with the following constraints: (i) due to the size of the latter, the evolution of $\mathfrak{s}$ has a negligible effect on that of $\mathfrak{S}$ and (ii) $\mathfrak{s}$ is robust in comparison to the strength of the interaction. By (ii) we mean that the characteristics of $\mathfrak{s}$ change *very* slowly, yet in such a way that the cumulative change over a *very* long time could be enormous.

At the conceptual level, the problem is to characterize the properties of $\mathfrak{s}$ over a (long) lapse of time, at least in non-trivial special cases. In particular, it is not always reasonable to expect that the subsystem $\mathfrak{s}$ reaches or remains in equilibrium under the perpetual influence of $\mathfrak{S}$. Rather, it might slowly pass through small vicinities of equilibria, in a continuous manner. Given an absence of equilibrium states, the question arises as to how to describe statistical properties of such a system.

In order to model the above set-up, we propose the notion of a *quasistatic dynamical system* (QDS) in §1.2. In §1.3 we introduce a means of describing properties of QDSs via stochastic diffusion processes. We follow this with a presentation of continuous-time QDSs and, subsequently, a discussion of relevant literature. In §2 we present a specific class of QDSs, whose properties we investigate in the remainder of the paper. Section 3 contains the statements of theorems.

1.2. *(Discrete-time) QDSs.* Throughout the paper, we use the standard notation

$$\{u\} = u \pmod 1 \quad \text{and} \quad \lfloor u \rfloor = u - \{u\}$$

for the fractional and integer part of a real number $u \geq 0$, respectively.

Definition 1.1. Let X be a set and $\mathcal{M}$ a collection of self-maps $T : X \to X$ equipped with a topology. Consider a triangular array

$$\mathbf{T} = \{T_{n,k} \in \mathcal{M} : 0 \leq k \leq n,\, n \geq 1\}$$

of elements of $\mathcal{M}$. If there exists a piecewise-continuous curve $\gamma : [0,\, 1] \to \mathcal{M}$ such that

$$\lim_{n\to\infty} T_{n,\lfloor nt \rfloor} = \gamma_t, \quad t \in [0,\, 1],$$

we say that $(\mathbf{T},\, \gamma)$ is a *quasistatic dynamical system (QDS).*

Definition 1.2. We call X the *state space* (also *phase space*) and $\mathcal{M}$ the *system space* of the QDS.

Obviously a QDS is a generalization of an ordinary dynamical system: a discrete-time dynamical system is determined by a single map $T : X \to X$, which is the degenerate QDS obtained by setting $T_{n,k} = T$ for all k and n. (Then $\gamma_t = T$ for all t.) Below, we will also provide the definition of a continuous-time QDS which similarly generalizes the notion of an ordinary continuous-time dynamical system (semiflow). One can readily come up with further generalizations of the definitions given, but spelling them out here seems to add little of substance for the purposes of this paper.

Let us justify the definition above. An element of the state space X represents the state of the system, whereas the triangular array $\mathbf{T}$ describes the dynamics: given the initial state $x \in X$ and integers $1 \leq k \leq n$, the image

$$x_{n,k} = T_{n,k} \circ \cdots \circ T_{n,1}(x) \in X$$

https://doi.org/10.1017/etds.2016.9 Published online by Cambridge University Press

is the state of the system after k steps on the nth level of the array $\mathbf{T}$. Passing to the continuous-time parameter $t \in [0, 1]$ via the scaling $k = \lfloor nt \rfloor$, we note that the piecewise-constant curve $t \mapsto T_{n,\lfloor nt \rfloor}$ approximates γ in the system space $\mathcal{M}$, and this approximation converges (pointwise) as $n \to \infty$. The system $T_{n,n} \approx \gamma_1$ differs significantly from the system $T_{n,1} \approx \gamma_0$ in typical situations.

The rate of convergence of $T_{n,\lfloor nt \rfloor}$ to γ_t will affect the properties of the QDS and may need to be specified when studying a particular system. Limiting properties will certainly depend on the limit curve γ and may or may not depend on the array $\mathbf{T}$ (see §3).

One cannot directly study a 'limit system', for letting first $n \to \infty$ one would end up just iterating the map γ_0 and none other. Rather, one must study (for example, statistical) properties of the QDS at level n and see how those properties behave in the limit. The main goal of this paper is to initiate a systematic study of such systems that would eventually result in a comprehensive mathematical theory of QDSs.

Drawing a parallel with the first paragraph of this section, the limit curve γ models the evolution of the observed system $\mathfrak{s}$, as the ambient system $\mathfrak{S}$ forces $\mathfrak{s}$ to transform, possibly significantly, over time. The jumps in γ allow for singular, abrupt events. For example, such events could include jumps from one connected component of the system space $\mathcal{M}$ to another.

We have borrowed the apt term 'quasistatic' from thermodynamics. There it refers to idealized processes in which the observed system transforms infinitesimally slowly due to external influence. Such a system is in thermodynamic equilibrium at any given time, yet traces out a continuous path of different equilibria over an infinitely long time span. For instance, all reversible thermodynamic processes are quasistatic (see, e.g., [**12**, **26**] for more background). The connection with thermodynamics should not be construed as a limitation of scope; the definition of QDS is, of course, purely abstract and could be used widely according to one's needs.

A natural class of QDS is provided by the following example.

Example 1.3. (Quasistatic billiards.) Dispersing billiards on a torus consists of strictly convex scatterers with smooth boundaries embedded in the (surface of the) torus. A particle moves on the torus, in the exterior of the scatterers, and experiences an elastic collision when it meets a scatterer. Assuming that the length of free flight between any two successive collisions is uniformly bounded, the location of the particle can be kept track of by keeping track of the collisions. For a fixed scatterer configuration K, this leads to a representation of the dynamics by a billiard map F_K mapping one collision to the next. A model of dispersing billiards with moving scatterers was introduced in [**33**]. Let $(K_k)_{k=0}^{\infty}$ be a sequence of scatterer configurations, such that $d(K_{k-1}, K_k) < \varepsilon$ holds uniformly for some small $\varepsilon > 0$. Here d is a natural distance on the space $\mathcal{K}$ of admissible configurations. Then $F_{K_k,K_{k-1}}$ represents the dynamics between the $(k-1)$th and kth collisions, during which the configuration has changed by a distance $<\varepsilon$, from K_{k-1} to K_k, and the compositions $F_{K_k,K_{k-1}} \circ \cdots \circ F_{K_1,K_0}$ represent the dynamics for scatterers moving with speed less than ε. This yields a QDS in the limit $\varepsilon \to 0$ of infinitesimally slowly moving scatterers. More precisely, consider a triangular array of configurations $\{K_{n,k} \in \mathcal{K} : 0 \le k \le n, n \ge 1\}$ with $d(K_{n,k-1}, K_{n,k}) < \varepsilon_n$, where

https://doi.org/10.1017/etds.2016.9 Published online by Cambridge University Press

$\lim_{n\to\infty} \varepsilon_n = 0$. Assume, moreover, that there is a continuous curve $\gamma^{\mathcal{K}} : [0, 1] \to \mathcal{K}$ such that $\lim_{n\to\infty} K_{n,\lfloor nt\rfloor} = \gamma_t^{\mathcal{K}}$. Setting $T_{n,k} = F_{K_{n,k},K_{n,k-1}}$ for $1 \le k \le n$ and $n \ge 1$ results in a QDS, which we coin quasistatic billiards. (The system space $\mathcal{M}$ can be identified with $\mathcal{K} \times \mathcal{K}$, in which case the limit curve γ takes values in the diagonal subspace: $\gamma_t = (\gamma_t^{\mathcal{K}}, \gamma_t^{\mathcal{K}})$.)

In this paper, we analyze a simpler model, introduced shortly in §2.

1.3. *Statistical description of a QDS.* In order to gather information about the state $x_{n,k}$ of the QDS, one performs a measurement of an observable quantity, whose value is determined by the state. Accordingly, the quantity is represented by a function $f : X \to \mathbb{R}$, called an 'observable', and the outcome of the measurement by $f(x_{n,k})$. We are going to study the sequence of measurements

$$f(x_{n,0}),\ f(x_{n,1}),\ \ldots,\ f(x_{n,n})$$

on the nth level of the QDS, corresponding to the initial state $x_{n,0} \equiv x$, as $n \to \infty$. For practical reasons, such as sensitive dependence on the initial state (also known as chaos), it is natural to assume that x is a random variable with values in X and some distribution μ. This renders $(x_{n,k})_{k=0}^n$ a sequence of random variables. By definition, each $x_{n,k}$ is completely determined by x, and its distribution is the pushforward† $\mu_{n,k} = (T_{n,k} \circ \cdots \circ T_{n,1})_*\mu$, which will be different for different n and k. In other words, the random variables $x_{n,k}$ are neither independent nor identically distributed.

The random paths

$$t \mapsto S_n(x, t) = \int_0^{nt} f(x_{n,\lfloor s\rfloor})\, ds$$

are piecewise-linear interpolations of Birkhoff sums: if $nt \in \mathbb{N}$, $S_n(x, t) = \sum_{j=0}^{nt-1} f(x_{n,j})$. Understanding the distribution of these paths for a large class of observables in the limit $n \to \infty$ is a natural goal and can be considered a good statistical description of a QDS. It is reasonable to hope that the random paths, once properly centered and scaled, converge to a stochastic diffusion process. We shall prove that this indeed happens for a paradigmatic QDS whose system space consists of strongly chaotic expanding circle maps, formally presented in §2. The result holds for rather generic f and μ; we only ask that f be Lipschitz continuous and μ absolutely continuous. The curve γ need not be especially regular; Hölder continuity suffices. The limit process will depend on γ and f, but not on μ (see §3 for the precise statements of the main results).

To get an idea of why passing to the limit $n \to \infty$ might yield a diffusion process for the paths, consider the following simple special case.

Example 1.4. (A degenerate case.) Let $T = T_{n,k} : x \mapsto 2x \pmod 1$ for all $n \ge 1$ and $0 \le k \le n$, so the curve γ is constant and $\gamma_t = T$ for all t. Define the observable $f : [0, 1) \to \mathbb{R}$ by $f(x) = -1$ for $x \in [0, \frac{1}{2})$ and $f(x) = 1$ otherwise. Taking Lebesgue measure for the initial distribution μ, the Birkhoff sums model a fair coin toss or the simple symmetric

† Assuming the map T is measurable, the pushforward of a measure μ is defined by $T_*\mu(E) = \mu(T^{-1}E)$ for all measurable sets E.

https://doi.org/10.1017/etds.2016.9 Published online by Cambridge University Press

random walk. The paths $n^{-1}S_n(x, \cdot)$ converge in distribution to the zero path $t \mapsto 0$. The random paths $n^{-1/2}S_n(x, \cdot)$ converge in distribution to standard Brownian motion.

1.4. *Continuous-time QDS.* For completeness, we finish this section with a discussion on continuous-time dynamical systems. The rest of the paper concerns a special discrete-time QDS introduced in §2. If the differential equation

$$\frac{dy}{ds} = V(y, s) \quad \text{with } y(r) = x$$

specified by the time-dependent vector field $V(\cdot, s)$ on X has a unique solution $y = y(x, r, s)$ for all $x \in X$ and all $r \geq 0$, then $\varphi^{r,s}(x) \equiv y(x, r, s)$ defines a time-dependent semiflow. (Given that the state x at time r, $\varphi^{r,s}(x)$ is the state s time units later.) Note that $\varphi^{r,s+u} = \varphi^{r+s,u} \circ \varphi^{r,s}$ with $\varphi^{r,0} = \mathrm{id}_X$.

Definition 1.5. Let X be a differentiable manifold and $\mathcal{M}$ a collection of vector fields on X equipped with a topology. Let n be a parameter taking values either in $\mathbb{Z}_+$ or $\mathbb{R}_+$. Consider the array

$$\mathbf{V} = \{V_n(\cdot, r) \in \mathcal{M} : 0 \leq r \leq n, n \geq 0\}$$

and assume that, for each $n \geq 0$, the time-dependent vector field $V_n(\cdot, r)$ determines a time-dependent semiflow $\varphi_n^{r,s}$, $0 \leq r \leq r + s \leq n$. If there exists a piecewise-continuous curve $\gamma : [0, 1] \to \mathcal{M}$ such that

$$\lim_{n\to\infty} V_n(\cdot, nt) = \gamma_t, \quad t \in [0, 1],$$

we say that $(\mathbf{V}, \gamma)$ is a *continuous-time QDS*.

Definition 1.6. The nomenclature of Definition 1.2 continues to apply in the continuous-time context. We also say that the array $\Phi = \{\varphi_n^{r,s} : 0 \leq r \leq r + s \leq n, n \geq 0\}$ is a *quasistatic semiflow*.

Again, one can readily come up with generalizations of the above definition (e.g. regarding the linear time scaling $r = nt$), but we do not record them here. Just as in the discrete-time case, it is interesting to study a quasistatic semiflow in terms of the limit behavior of the random paths $t \mapsto S_n(x, t) = \int_0^{nt} f(x_{n,s})\, ds$, where $x_{n,s} = \varphi_n^{0,s}(x)$.

1.5. *Preceding literature.* This work is philosophically a natural successor of the sequence of papers [**20**, **29**, **31**, **33**] on the analysis of time-dependent dynamical systems, although in scope it differs drastically. In [**20**], 'asymptotic similarity' of pushforward densities under sequences of (piecewise) expanding maps was established using transfer (Perron–Frobenius) operator techniques. Asymptotic similarity is an indication of statistical *memory loss*: the system quickly forgets the distribution of its initial state. In [**29**, **31**, **33**], a different approach for establishing memory loss for progressively more complicated time-dependent systems was developed based on the idea of *coupling*. (See [**6–8**, **35**] for implementations of coupling in deterministic dynamics and [**23**] for an introduction to the probability theory of coupling). Let us also mention [**15**, **27**], which use yet another technique (Hilbert projective metric) for a time-dependent system, as well as [**1**]. The source of memory loss in the systems of the preceding papers is sensitive

https://doi.org/10.1017/etds.2016.9 Published online by Cambridge University Press

dependence on initial conditions (also known as chaos). In the other extreme, memory loss is also produced by sinks; the references closest to the time-dependent set-up we have been able to find concern *random sinks* [**4**, **21**]. In addition to memory loss issues, there are works including [**16**, **18**, **19**, **36**] on the entropy of time-dependent dynamical systems.

Besides the ones mentioned, few limit laws have been proved on the properties of non-random time-dependent dynamical systems. (On the contrary, a vast literature—which we cannot even begin to cover here—exists on random dynamical systems, concerning both averaged/annealed and quenched limit theorems.) In special cases, some central limit theorems have been obtained [**2**, **9**, **28**]. Indeed, one of the difficulties one faces in this setting is that it is often not even clear how possible results should be formulated, let alone proved.

In our proof, a key ingredient is rapid memory loss, which we establish via coupling. In order to identify the limit process, we solve a well-posed martigale problem [**13**, **30**, **34**]. In the context of dynamical systems, the idea of resolving to a martingale problem has been used (sparingly) in the theory of averaging: first, to our knowledge, in [**11**], and then in [**10**], which is closest to our work. Averaging is a tool in the analysis of so-called slow–fast systems concerning the limit where one of two variables evolves infinitely fast compared to the other: in order to describe the evolution of the *slow* variable, one may be able to 'average out' the influence of the fast one and thus reduce the problem to an effective one in which the slow variable alone appears. We stress that the abstract set-up in §1.1 is situated rather at the other end of the spectrum: it involves studying the fast variable $x_{n,k}$, which (in the limit $n \to \infty$) describes the state of the observed system $\mathfrak{s}$ under the influence of the large, slow, system $\mathfrak{S}$. Let us mention, however, that averaging ideas could be used to study the statistical properties of QDSs in certain situations, regarding the normalized Birkhoff sum $n^{-1/2} S_n(x, t)$ (along with an index) as a slow variable and x as a fast variable.

1.6. *How the paper is organized.* The model QDS, considered in the rest of the paper, is presented in §2. In §3, we state our two main results, Theorems 3.1 and 3.6, keeping the technical prerequisites down to a minimum. In §4, we explain frequently used notation and record key definitions. We then introduce, in §§5 and 6, the preliminaries necessary for understanding the proofs of Theorems 3.1 and 3.6, which are presented in §§7 and 8, respectively. In §3.3, we also state generalizations of the main results, whose proofs are outlined in §9; the proofs involve minor modifications of the preceding sections, so we only indicate the necessary changes, with the hope that such a choice makes the paper easier to read.

2. *The model*

In this section, we introduce the QDS to be studied in the rest of the paper. Fix $\lambda > 1$ and $A_* > 0$ once and for all. Let $\mathcal{M}$ denote the set of C^2 expanding maps $T : \mathbb{S} \to \mathbb{S}$ on the circle with the bounds†

$$\inf T' \geq \lambda \quad \text{and} \quad \|T''\|_\infty \leq A_*, \qquad T \in \mathcal{M}.$$

† Throughout, we use notation such as $\inf T'$ for $\inf_{x \in \mathbb{S}} T'(x)$.

https://doi.org/10.1017/etds.2016.9 Published online by Cambridge University Press

The space $\mathcal{M}$ is endowed with the metric d_{C^1} defined by

$$d_{C^1}(T_1, T_2) = \sup_{x\in\mathbb{S}} d(T_1x, T_2x) + \|T_1' - T_2'\|_\infty$$

for $T_1, T_2 \in \mathcal{M}$. Here d is the natural metric on $\mathbb{S} = \mathbb{R}/\mathbb{Z}$.

We construct a QDS with state space $\mathbb{S}$ and system space $\mathcal{M}$, as follows. First, fix a Hölder continuous curve $\gamma : [0, 1] \to \mathcal{M}$ with exponent $\eta \in (0, 1)$. Let $\mathbf{T}$ be a triangular array of maps

$$\mathbf{T} = \{T_{n,k} \in \mathcal{M} : 0 \le k \le n, n \ge 1\}$$

for which

$$\sup_{n\ge1} n^\eta \sup_{0\le t\le1} d_{C^1}(T_{n,\lfloor nt\rfloor}, \gamma_t) < \infty. \tag{1}$$

Clearly, $(\mathbf{T}, \gamma)$ meets the requirements of Definition 1.1 of a QDS. A prototypical example to keep in mind is where $T_{n,k} = \gamma_{kn^{-1}}$, although the maps $T_{n,k}$ are not required to live on the curve γ. The convergence rate (1) is chosen to reflect the smoothness of the curve; in particular, we have similar bounds on $d_{C^1}(T_{n,k}, \gamma_{kn^{-1}})$ and on $d_{C^1}(\gamma_{kn^{-1}}, \gamma_{(k+1)n^{-1}})$, resulting in the bound

$$\sup_{n\ge1} n^\eta \sup_{0\le k<n} d_{C^1}(T_{n,k}, T_{n,k+1}) < \infty.$$

Stronger convergence would not be a natural assumption.

For a fixed n, the maps $T_{n,k}$ approximate the curve γ, traversing from beginning to end as k increases from 1 to n. Moreover, $T_{n,\lfloor nt\rfloor}$ tends to the well-defined limit γ_t as $n \to \infty$.

We choose to work with circle maps as these are continuous maps with the same domain of definition; they lie in discrete topological classes, so maps along a continuous path are all topologically conjugate. This simplifies definitions and subtleties concerning points of discontinuity, for example, allowing us to focus on the heart of the issue. For future use, we point out that every $T \in \mathcal{M}$ has a unique invariant probability measure $\hat{\mu}_T$ equivalent to the Lebesgue measure $\mathfrak{m}$ on $\mathbb{S}$ [**3**]. This is a consequence of expansion. The measure $\hat{\mu}_T$ we sometimes call an *SRB measure*, for Sinai–Ruelle–Bowen. For brevity, we write

$$\hat{\mu}_t = \hat{\mu}_{\gamma_t} \quad \text{and} \quad \hat{\mu}_{n,k} = \hat{\mu}_{T_{n,k}}.$$

3. *Results*

We begin our study of the QDS $(\mathbf{T}, \gamma)$ introduced in §2. The goal is to understand statistical properties of the QDS. So let $f : \mathbb{S} \to \mathbb{R}$ be an observable, and denote

$$f_{n,k} = f \circ T_{n,k} \circ \cdots \circ T_{n,1}, \quad 0 \le k \le n. \tag{2}$$

(Our convention is that $f_{n,0} = f$.) This yields a triangular array of random variables once an initial distribution μ on $\mathbb{S}$ is given. We define the functions $S_n : \mathbb{S} \times [0, 1] \to \mathbb{R}$ by

$$S_n(x, t) = \int_0^{nt} f_{n,\lfloor s\rfloor}(x)\, ds = \sum_{k=0}^{\lfloor nt\rfloor-1} f_{n,k}(x) + \{nt\} f_{n,\lfloor nt\rfloor}(x), \quad n \ge 1.$$

https://doi.org/10.1017/etds.2016.9 Published online by Cambridge University Press

3.1. *The mean.* A natural quantity to study first is the mean

$$\zeta_n(x, t) = n^{-1} S_n(x, t). \tag{3}$$

Given an initial probability measure μ on $\mathbb{S}$, each ζ_n is a random element of $C^0([0, 1], \mathbb{R})$, whose distribution we denote $\mathbf{P}_n^\mu$. As is customary, we also denote $\hat{\mu}_t(f) = \int f \, d\hat{\mu}_t$, and so on.

THEOREM 3.1. *Suppose f is Lipschitz continuous and μ is absolutely continuous. The function $t \mapsto \hat{\mu}_t(f)$ is continuous. The measures $\mathbf{P}_n^\mu$ converge weakly, as $n \to \infty$, to the point mass at $\zeta \in C^0([0, 1], \mathbb{R})$, where*

$$\zeta(t) = \int_0^t \hat{\mu}_s(f) \, ds. \tag{4}$$

Remark 3.2. Note that the limit distribution in Theorem 3.1 is independent of μ.

In other words, given an arbitrary initial measure μ having a density, the stochastic process ζ_n converges, as $n \to \infty$, to the non-random limit ζ.

3.2. *Fluctuations about the mean.* The mean gives a coarse description of a limit statistical property. Deeper insight is obtained by studying the fluctuations at a finer scale, by looking into the statistical properties of $n^{1/2}\zeta_n$, instead of ζ_n. For this to make sense, a centering is needed. To this end, we choose a sequence of functions $c_n : \mathbb{S} \times [0, 1] \to \mathbb{R}$, and define the functions $\chi_n : \mathbb{S} \times [0, 1] \to \mathbb{R}$ by

$$\chi_n(x, t) = n^{1/2}\zeta_n(x, t) - n^{1/2}c_n(x, t), \tag{5}$$

quantifying the fluctuations of $n^{1/2}\zeta_n$ about $n^{1/2}c_n$. The goal is to describe the statistics of these fluctuations, as $n \to \infty$, by a probabilistic limit law.

The choice of a good centering sequence $(c_n)_{n\geq 1}$ turns out to be a delicate issue. There are two canonical choices, both independent of x, namely, $c_n(t) = \mu(\zeta_n(\,\cdot\,, t))$ (where μ is the initial measure) and $c_n(t) = \mathfrak{m}(\zeta_n(\,\cdot\,, t))$ (where $\mathfrak{m}$ is the Lebesgue measure). The first of these amounts to centering by force: $\mu(\chi_n(\,\cdot\,, t)) = 0$ holds for all $t \in [0, 1]$. It is perhaps surprising that this 'natural' choice does *not* appear to yield good statistics for χ_n unless the initial measure μ has a sufficiently regular density. The second choice, centering using the Lebesgue measure instead, works better: in that case we obtain a universal limit law for χ_n as long as the initial measure μ is just *absolutely continuous*. We now introduce the notion of an admissible centering sequence, for which our main result holds.

Definition 3.3. We say that a centering sequence $(c_n)_{n\geq 1}$ of functions $c_n : \mathbb{S} \times [0, 1] \to \mathbb{R}$ is *admissible* with respect to an initial probability measure μ, if (i) $c_n(x, \cdot\,)$ is continuous for almost every x with respect to μ and (ii) $t \mapsto n^{1/2}c_n(x, t) - n^{1/2}\mathfrak{m}(\zeta_n(\,\cdot\,, t))$ viewed as a random element of $C^0([0, 1], \mathbb{R})$ converges to zero in probability with respect to μ: that is,

$$\lim_{n\to\infty} \mu\Big(\Big\{x \in \mathbb{S} : \sup_{t\in[0,1]} |n^{1/2}c_n(x, t) - n^{1/2}\mathfrak{m}(\zeta_n(\,\cdot\,, t))| > \delta\Big\}\Big) = 0$$

for any $\delta > 0$. For brevity, we also say, in this case, that c_n is admissible (with respect to μ).

https://doi.org/10.1017/etds.2016.9 Published online by Cambridge University Press

Remark 3.4. If the centering c_n is independent of x, there is no need to refer to an initial measure: c_n is admissible either with respect to every measure or no measure at all; it is admissible if and only if

$$\lim_{n\to\infty} \sup_{t\in[0,1]} |n^{1/2}c_n(t) - n^{1/2}\mathfrak{m}(\zeta_n(\,\cdot\,,t))| = 0. \tag{6}$$

Clearly, $c_n(t) = \mathfrak{m}(\zeta_n(\,\cdot\,,t))$ is admissible. Since $\mathfrak{m}(\zeta_n(\,\cdot\,,t))$ converges to $\zeta(t)$, so does $c_n(t)$ for any admissible c_n. Whether $n^{1/2}\mathfrak{m}(\zeta_n(\,\cdot\,,t))$ converges to $n^{1/2}\zeta(t)$ (i.e. whether $\zeta(t)$ is admissible) seems to depend on the regularity of the curve γ (see Lemma 3.5(ii)). On the other hand, given a measure ν (perhaps different from the initial measure μ), $c_n(t) = \nu(\zeta_n(\,\cdot\,,t))$ may or may not be admissible, depending on the regularity of (the density of) ν.

The next lemma establishes admissible centering sequences of the preceding kind under different conditions, and sheds light on the role of the function ζ defined in (4).

LEMMA 3.5.

(i) *If $\eta \in (0,1)$ is arbitrary and ν is an arbitrary probability measure having a Lipschitz continuous density, then the centering $c_n(t) = \nu(\zeta_n(\,\cdot\,,t))$ is admissible.*

(ii) *If $\eta > \frac{1}{2}$, then the* explicit centering

$$c_n(t) = \zeta(t)$$

is admissible. Here ζ is the function defined in (4).

(iii) *If $\eta \in (0,1)$ is arbitrary and ν is an arbitrary absolutely continuous measure, then $\nu(\zeta_n(\,\cdot\,,t))$ tends to $\zeta(t)$, as $n \to \infty$. The convergence is uniform in t.*

Lemma 3.5 is proved in §8.1.

We are now ready to proceed to the main result of the paper. Given an initial probability measure μ on $\mathbb{S}$ and an admissible centering sequence $(c_n)_{n\geq 1}$, each χ_n is a random element of $C^0([0,1],\mathbb{R})$, whose distribution we denote $\mathbb{P}_n^\mu$. We also denote

$$\hat{f}_t = f - \hat{\mu}_t(f)$$

and

$$\hat{\sigma}_t^2(f) = \lim_{m\to\infty} \hat{\mu}_t\left[\left(\frac{1}{\sqrt{m}}\sum_{k=0}^{m-1} \hat{f}_t \circ \gamma_t^k\right)^2\right], \tag{7}$$

which is just the limit of the variance of $(1/\sqrt{m})\sum_{k=0}^{m-1} f\circ\gamma_t^k$ with respect to the measure $\hat{\mu}_t$.

THEOREM 3.6. *Suppose f is Lipschitz continuous, μ is absolutely continuous and $(c_n)_{n\geq 0}$ is admissible. The function $t \mapsto \hat{\sigma}_t^2(f)$ is continuous. The measures $\mathbb{P}_n^\mu$ converge weakly, as $n \to \infty$, to the law of the process*

$$\chi(t) = \int_0^t \hat{\sigma}_s(f)\, dW_s.$$

Here W is a standard Brownian motion and the stochastic integral is to be understood in the sense of Itō.

https://doi.org/10.1017/etds.2016.9 Published online by Cambridge University Press

Let us pause to discuss Theorem 3.6 in conjunction with Lemma 3.5. Here ν is an arbitrary measure having a Lipschitz continuous density. For every value of the regularity exponent $\eta \in (0, 1)$ of the curve γ, the *implicit* centering $c_n(t) = \nu(\zeta_n(\,\cdot\,, t))$ is a valid choice; in particular, $c_n(t) = \mathfrak{m}(\zeta_n(\,\cdot\,, t))$ can always be used, and, if the initial measure μ has a Lipschitz continuous density (some hypothesis on the density is necessary; see Example 3.7), the choice $c_n(t) = \mu(\zeta_n(\,\cdot\,, t))$ is also admissible. The special case $\eta < \frac{1}{2}$ is interesting from the point of view that γ is allowed to have the degree of regularity of Brownian paths. (With probability one, the path of a Brownian motion is Hölder continuous with any exponent $< \frac{1}{2}$.) Thus, we may think, in this case, that the driving ambient system forces the observed system γ_t to perform a Brownian motion in the space $\mathcal{M}$. On the other hand, the case $\eta > \frac{1}{2}$ of a more regular curve γ allows also for the *explicit centering* $c_n(t) = \zeta(t)$ to be used. The reason for this dichotomy is that the order of $n^{1/2}\nu(\zeta_n(\,\cdot\,, t)) - n^{1/2}\zeta(t)$ appears to be large enough to contribute to the limit process, if the regularity exponent η of the curve γ is $\frac{1}{2}$ or less. A similar remark concerns the order of $n^{1/2}\mu(\zeta_n(\,\cdot\,, t)) - n^{1/2}\nu(\zeta_n(\,\cdot\,, t))$, if the density of μ is not sufficiently regular. Let us illustrate the last point with an example.

Example 3.7. (An inadmissible centering.) Let us return to the setting of Example 1.4, so $T = T_{n,k} : x \mapsto 2x \pmod 1$ for all n, k, $f(x) = -1$ for $x \in [0, \frac{1}{2})$ and $f(x) = +1$ otherwise. Then $\mathfrak{m}(\zeta_n(\,\cdot\,, t)) = 0$ for all n, t. Taking as our centering the zero function, the random paths $\chi_n(\,\cdot\,)$ converge in distribution to standard Brownian motion, if Lebesgue measure is the initial distribution. Disregarding the (here inconsequential) non-Lipschitz nature of f, Theorem 3.6 states† that, for *any* absolutely continuous initial distribution, the random paths converge in distribution to the same process—standard Brownian motion. This need not hold if one changes the centering.

For $j \geq 1$, let A_j denote the interval $[0, 2^{-2^j})$ and define the probability measure $\nu_j = 2^{2^j} 1_{A_j}\mathfrak{m}$. For $k \leq 2^j - 1$, $T^k(A_j) \subset [0, \frac{1}{2})$ so $\nu_j(f \circ T^k) = -1$, while, for $k \geq 2^j$, the pushforward $T^k_* \nu_j = \mathfrak{m}$ so $\nu_j(f \circ T^k) = \mathfrak{m}(f) = 0$. Therefore $\nu_j(\zeta_n(\,\cdot\,, t)) \leq 0$ for all j, n, t. Let $K \in \mathbb{N}$ be large enough so that $\varepsilon = \sum_{j \geq K} (1/j^2) < \frac{1}{2}$. Let μ be the absolutely continuous probability measure

$$\mu = (1 - \varepsilon)\mathfrak{m} + \sum_{j=K}^{\infty} \frac{1}{j^2} \nu_j.$$

Since $\nu_j(\zeta_n(\,\cdot\,, t)) \leq 0$ for each j,

$$\mu(\zeta_n(\,\cdot\,, t)) \leq \frac{1}{j^2} \nu_j(\zeta_n(\,\cdot\,, t)).$$

Let $t \in (0, 1]$ and take n big enough so that $nt > 2^K$. Denote $j_n = \lfloor \log_2 \lfloor nt \rfloor \rfloor$, so $2^{j_n} \leq \lfloor nt \rfloor < 2^{j_n+1}$. It follows that $K \leq j_n \leq \log_2 nt$. Now $\nu_{j_n+1}(\zeta_n(\,\cdot\,, t)) = n^{-1} \int_0^{nt} (-1)\, ds = -t$. Therefore,

$$n^{1/2}\mu(\zeta_n(\,\cdot\,, t)) \leq n^{1/2} \frac{1}{(j_n + 1)^2} \nu_{j_n+1}(\zeta_n(\,\cdot\,, t)) = -\frac{tn^{1/2}}{(j_n + 1)^2} \leq -\frac{tn^{1/2}}{(\log_2 nt + 1)^2},$$

† Theorem 3.6 can be extended to the set-up of the example. Alternatively, a Lipschitz approximation of f can be used to obtain similar results; we skip the straightforward details.

https://doi.org/10.1017/etds.2016.9 Published online by Cambridge University Press

which tends to $-\infty$ as $n \to \infty$. In particular, if one sets $c_n(t) = \mu(\zeta_n(\,\cdot\,, t))$, then $n^{1/2}c_n$ converges pointwise (and uniformly away from zero) to $-\infty$, while $n^{1/2}\mathfrak{m}(\zeta_n(\,\cdot\,, t)) = 0$. Hence, $(c_n)_{n\geq 1}$ does not form an admissible centering sequence (see (6)).

Remark 3.8.

(i) Note that the limit distribution in Theorem 3.6 is independent of μ.

(ii) The process χ is a martingale starting at zero and its quadratic variation is

$$[\chi]_t = \int_0^t \hat{\sigma}_s^2(f)\, ds.$$

It follows [**30**] that there exists such a standard Brownian motion $\widetilde{W}$ that

$$\chi(t) = \widetilde{W}_{[\chi]_t}.$$

In other words, χ is a Brownian motion up to the non-random time change $t \mapsto [\chi]_t$.

(iii) Regarding the possible degeneracy of the process χ, a result due to Leonov [**22**] shows that, given $s \in [0, 1]$, $\hat{\sigma}_s^2(f) = 0$ if and only if there exists an L^2 function $g : \mathbb{S} \to \mathbb{R}$ such that the cohomology equation $\hat{f}_s = g \circ \gamma_s - g$ holds almost everywhere, and Livschitz (Livšic) rigidity theory [**24**, **25**] shows that g has a Hölder continuous representative (with the same exponent as f) for which the cohomology equation holds everywhere. Equivalently, $\sum_{k=0}^{p-1} \hat{f}_s \circ \gamma_s^k(x) = 0$ for any point x of any period p with respect to γ_s. Since periodic points are dense, $\hat{\sigma}_s^2(f) = 0$ amounts to a severe restriction on the choice of f. Non-degeneracy is, moreover, an open condition with respect to the map: if $\hat{\sigma}_s^2(f) \neq 0$, then $\hat{\sigma}_t^2(f) \neq 0$ for all t sufficiently close to s (see Lemma 6.1).

3.3. *Generalizations.* Here we discuss two ways to generalize Theorems 3.1 and 3.6. First, we consider vector-valued observables $f : \mathbb{S} \to \mathbb{R}^d$, $d > 1$. After that, we consider curves γ in $\mathcal{M}$ with a finite number of discontinuities (jumps).

In the case of a vector-valued observable f, we define the vector-valued processes ζ_n, ζ and χ_n according to the formulas (3), (4) and (5), respectively. The centering c_n is also vector-valued, but Definition 3.3 remains otherwise intact. For clarity, we denote the vector components by superindices enclosed in parentheses: $f = (f^{(1)}, \ldots, f^{(d)})$, and so on.

THEOREM 3.9. *Suppose $f : \mathbb{S} \to \mathbb{R}^d$, $d > 1$, is Lipschitz continuous. Then Theorems 3.1 and 3.6 as well as Lemma 3.5 continue to hold with the modifications that W is a d-dimensional standard Brownian motion and $\hat{\sigma}_t(f)$ is the $d \times d$ matrix defined as the square root of the covariance matrix*

$$\hat{\sigma}_t^2(f) = \lim_{m\to\infty} \hat{\mu}_t\left[\left(\frac{1}{\sqrt{m}}\sum_{k=0}^{m-1} \hat{f}_t \circ \gamma_t^k\right) \otimes \left(\frac{1}{\sqrt{m}}\sum_{k=0}^{m-1} \hat{f}_t \circ \gamma_t^k\right)\right].$$

Here $(v \otimes v)_{ij} = v^{(i)}v^{(j)}$ for $v = (v^{(1)}, \ldots, v^{(d)}) \in \mathbb{R}^d$.

Note that $\zeta_n^{(i)}$ and $\chi_n^{(i)}$ are exactly the processes appearing in (3) and (5) for the scalar-valued observable $f^{(i)}$, $1 \leq i \leq d$. The generalization here is that Theorem 3.9 describes the joint behavior of each of these two sets of d scalar-valued processes.

The next result allows for jump discontinuities.

https://doi.org/10.1017/etds.2016.9 Published online by Cambridge University Press

THEOREM 3.10. *Let* $\{I_1, \ldots, I_m\}$ *be a finite partition of* $[0, 1]$ *into intervals. Suppose the curve* $\gamma : [0, 1] \to \mathcal{M}$ *restricted to each* I_i, $1 \le i \le m$, *is Hölder continuous with exponent* $\eta \in (0, 1)$, *having possibly jumps where two intervals meet. Then Theorems 3.1 and 3.6 as well as Lemma 3.5 continue to hold, as does Theorem 3.9, with the modifications that the functions* $t \mapsto \hat{\mu}_t(f)$ *and* $t \mapsto \hat{\sigma}_t^2(f)$ *are piecewise continuous.*

In particular, the curve γ may pass through different connected components of the system space $\mathcal{M}$. For instance, the number of inverse branches of the map γ_t does not have to remain constant with t.

4. *Notation and definitions*

We use Lip to denote the space of Lipschitz continuous maps (from $\mathbb{S}$ to $\mathbb{R}$), $\mathrm{Lip}(h)$ to denote the Lipschitz constant of a function h and $\|h\|_{\mathrm{Lip}} = \mathrm{Lip}(h) + \|h\|_\infty$ to denote its Lipschitz norm.

Given $T \in \mathcal{M}$, the transfer operator $\mathcal{L}_T : L^1(\mathfrak{m}) \to L^1(\mathfrak{m})$ is defined by

$$\mathcal{L}_T h(x) = \sum_{y \in T^{-1}\{x\}} \frac{h(y)}{T'(y)}.$$

The basic duality property of $\mathcal{L}_T$, which we use repeatedly, is encoded in the identity

$$\mathfrak{m}(f \circ T \cdot h) = \mathfrak{m}(f \cdot \mathcal{L}_T h)$$

for $f \in L^\infty(\mathfrak{m})$ and $h \in L^1(\mathfrak{m})$. The transfer operator describes the evolution of probability densities under the map T: if x is distributed according to probability density ρ, then $T(x)$ is distributed according to the probability density $\mathcal{L}_T \rho$. This (Perron–Frobenius) transfer operator is positive with a unique maximal eigenvalue 1. Its corresponding (and unique positive) eigenvector is the density of the SRB for T (see [**3**] for details). We will often write

$$\mathcal{L}_t = \mathcal{L}_{\gamma_t} \quad \text{and} \quad \mathcal{L}_{n,k} = \mathcal{L}_{T_{n,k}}.$$

If $\mathcal{L}_i$ is the transfer operator of T_i, then the transfer operator of $\mathcal{T} = T_k \circ \cdots \circ T_1$ satisfies $\mathcal{L}_\mathcal{T} = \mathcal{L}_k \cdots \mathcal{L}_1$. We shall sometimes write $\mathcal{T}z$ for $\mathcal{T}(z)$.

In the following section, $\mathcal{L}$ will stand for the transfer operator of a general map in $\mathcal{M}$ and $(\mathcal{L}_i)_{i=1}^\infty$ will be a general sequence of such operators. The constants appearing will not depend on the choice of these operators. We shall routinely use the facts that $\mathcal{L}$ maps probability densities to probability densities and that

$$|\mathcal{L}(uv)| \le \mathcal{L}|uv| \le \|u\|_\infty \mathcal{L}|v|.$$

Recall that $\hat{\mu}_t$ is the SRB measure (equivalent to Lebesgue measure $\mathfrak{m}$) associated to γ_t and that $\hat{\mu}_{n,k} \equiv \hat{\mu}_{T_{n,k}}$ is the measure associated to $T_{n,k}$. Denote the corresponding densities by $\hat{\rho}_t$ and $\hat{\rho}_{n,k}$ and the density of $\hat{\mu}_T$ by $\hat{\rho}_T$.

Remark 4.1. Since $\mathcal{L}_T^k 1$ converges in the supremum norm to $\hat{\rho}_T$ as $k \to \infty$, it will follow, from (12), that $\hat{\rho}_T \in \mathrm{Lip}$ with

$$\sup_{T \in \mathcal{M}} \|\hat{\rho}_T\|_{\mathrm{Lip}} < \infty. \tag{8}$$

https://doi.org/10.1017/etds.2016.9 Published online by Cambridge University Press

Let $\mu_{n,k}$ denote the pushforward

$$\mu_{n,k} = (T_{n,k} \circ \cdots \circ T_{n,1})_* \mu,$$

where μ is a probability measure with density ρ. Then the density of $\mu_{n,k}$ is

$$\rho_{n,k} = \mathcal{L}_{n,k} \cdots \mathcal{L}_{n,1} \rho.$$

Recall that $f_{n,k} = f \circ T_{n,k} \circ \cdots \circ T_{n,1}$ (see (2)). Given an initial measure μ, we will often encounter the centered versions of such functions, defined as

$$\bar{f}_{n,k} = f_{n,k} - \mu(f_{n,k}). \tag{9}$$

Our main results are formulated in terms of Lipschitz continuous densities on $\mathbb{S}$. However, for technical reasons, we introduce the following notions of regularity which will be convenient for the proofs.

Definition 4.2. Given a point $z \in \mathbb{S}$, we define the arc $J_z = \mathbb{S} \backslash \{z\}$. We denote by $|x - y|_z$ the length of the subarc of J_z with endpoints x, y. We say that a function $g : J_z \to \mathbb{R}$ is Lipschitz continuous on J_z, with constant $L > 0$, if

$$|g(x) - g(y)| \le L|x - y|_z, \quad x, y \in J_z.$$

In other words, g is Lipschitz continuous except across $z \in \mathbb{S}$ when the distance of x and y is understood as the length of the arc between the two points not containing z.

Definition 4.3. Given $L > 0$, let $\mathcal{D}_L$ be the class of all probability densities $\psi : \mathbb{S} \to \mathbb{R}$ with the properties that (i) $\psi > 0$ and (ii) there exists $z \in \mathbb{S}$ such that $\log \psi$ is Lipschitz continuous on $J_z = \mathbb{S} \backslash \{z\}$ with constant L.

The motivation for these definitions is that we will need to study conditional probability densities of the form $\psi_I = |\int_I \psi \, d\mathrm{m}|^{-1} \psi 1_I$, where $I \subset \mathbb{S}$ is a sufficiently short arc, and their pushforwards $\mathcal{L}_{\mathcal{T}}(\psi_I)$ under compositions $\mathcal{T} = T_k \circ \cdots \circ T_1$. Such pushforwards are generally discontinuous on $\mathbb{S}$, even if ψ is continuous, and repeating the procedure can lead to a growing number of discontinuities. Being more careful, this nuisance can be circumvented. If ψ is continuous, except possibly at the single point $z \in \mathbb{S}$, we may choose the intervals I to be the partition elements induced by the points $\mathcal{T}^{-1}\mathcal{T}z$. Then $\mathcal{L}_{\mathcal{T}}(\psi_I)$ is continuous, except, possibly, at the single point $\mathcal{T}z$. The details are discussed in §5.

For the most part, we will work with the class $\mathcal{D}_L$ and prove the main theorems in the setting of such initial densities first. Only then will the theorems be generalized to include Lipschitz continuous initial densities which are allowed to take the value zero.

Remark 4.4.

(i) A function $g : \mathbb{S} \to \mathbb{R}$ that is Lipschitz continuous with constant L relative to the standard metric d of $\mathbb{S}$ is, in the above sense, Lipschitz continuous on J_z with the same constant L, for any $z \in \mathbb{S}$. This follows from $d(x, y) \le |x - y|_z$.

(ii) Note that if $\psi \in \mathcal{D}_L$, then $e^{-L} \le \psi \le e^L$ and that ψ itself (without the logarithm) is Lipschitz continuous on J_z with constant Le^L. These follow from the fact that the probability density ψ must take the value 1 somewhere on J_z.

https://doi.org/10.1017/etds.2016.9 Published online by Cambridge University Press

(iii) In the opposite direction, if a probability density ψ satisfies $\psi \geq c > 0$ and is Lipschitz continuous on J_z with constant L, then $\psi \in \mathcal{D}_{c^{-1}L}$. Indeed,

$$|\log \psi(x) - \log \psi(y)| \leq c^{-1}|\psi(x) - \psi(y)| \leq c^{-1}L|x - y|_z, \quad x, y \in J_z.$$

The following lemma states that $\bigcup_{L>0} \mathcal{D}_L$ is dense in the set of all probability densities with respect to the L^1 norm. We leave the proof as an exercise.

LEMMA 4.5. *Let $\varphi : \mathbb{S} \to \mathbb{R}$ be an arbitrary probability density. Given any $\varepsilon > 0$, there exists $\psi \in \bigcup_{L>0} \mathcal{D}_L$ such that*

$$\|\varphi - \psi\|_{L^1} \leq \varepsilon.$$

5. *Preliminaries I: densities, composition and decorrelation*

The conceptually important results of this section are: Lemma 5.2, which gives uniform control of the Lipschitz constants for the densities of pushforwards; Lemma 5.9, which says that the pushforward density $\rho_{n,\lfloor nt \rfloor}$ is pretty close to the SRB density $\hat{\rho}_s$, provided the system has been running for a while and provided s is close to t; and the exponential decorrelation results of Corollary 5.7 and Lemma 5.10.

Recall that $1 < \lambda \leq T'$ and $\|T''\|_\infty \leq A_* < \infty$ for all maps $T \in \mathcal{M}$.

LEMMA 5.1. *Denoting by $\mathcal{T}$ any composition of a finite sequence of maps from $\mathcal{M}$,*

$$\sup_{\mathcal{T}} \sup_{x \in \mathbb{S}} \left| \frac{\mathcal{T}''(x)}{\mathcal{T}'(x)^2} \right| < \infty. \tag{10}$$

Moreover, there exists $C > 0$ such that, for all such $\mathcal{T}$, if I is an arc on which $\mathcal{T}$ is diffeomorphic,

$$\left| \frac{\mathcal{T}'(x)}{\mathcal{T}'(y)} - 1 \right| \leq C\, |\mathcal{T}(I_{x,y})| \tag{11}$$

for all $x, y \in I$, where $I_{x,y}$ denotes the subarc of I connecting x and y.

Proof. Suppose $\mathcal{T} = T_k \circ \cdots \circ T_1$, where $T_j \in \mathcal{M}$. For $j \geq 2$, write $P_j = T_{j-1} \circ \cdots \circ T_1$ and let P_1 be the identity map. Then

$$\mathcal{T}''(x) = \mathcal{T}'(x) \sum_{j=1}^{k} \frac{T_j''(P_j(x))}{T_j'(P_j(x))} P_j'(x),$$

so, noting $|P_j'(x)/\mathcal{T}'(x)| \leq \lambda^{j-1-k}$,

$$\left| \frac{\mathcal{T}''(x)}{\mathcal{T}'(x)^2} \right| \leq A_* \sum_{j=0}^{k-1} \lambda^{j-k},$$

from which (10) follows.

By the change-of-variables formula and (10),

$$\left| \log \frac{\mathcal{T}'(x)}{\mathcal{T}'(y)} \right| = \left| \int_{I_{x,y}} \frac{\mathcal{T}''}{\mathcal{T}'}\, d\mathrm{m} \right| \leq \int_{\mathcal{T}(I_{x,y})} \left\| \frac{\mathcal{T}''}{(\mathcal{T}')^2} \right\|_\infty d\mathrm{m} \leq C\, |\mathcal{T}(I_{x,y})|.$$

Now (11) follows from the estimate $|u - 1| \leq e^{|\log u|} |\log u|$, $u > 0$. □

https://doi.org/10.1017/etds.2016.9 Published online by Cambridge University Press

LEMMA 5.2. *There exists $L_* > 1$ such that the following hold.*

(i) *For all $L \geq L_*$ and $T \in \mathcal{M}$,*

$$\mathcal{L}_T \mathcal{D}_L \subset \mathcal{D}_L.$$

(ii) *Given $L > L_*$, there exists a constant $N(L)$ such that*

$$\mathcal{L}_k \cdots \mathcal{L}_1 \mathcal{D}_L \subset \mathcal{D}_{L_*}, \quad k \geq N(L),$$

for any sequence of maps $(T_i) \subset \mathcal{M}$.

(iii) *Let $L \geq L_*$ and $z \in \mathbb{S}$. Suppose $\psi \in \mathcal{D}_L$ with $|\log \psi(x) - \log \psi(y)| \leq L|x-y|_z$ for $x, y \in J_z$. Let $\mathcal{T}$ be a composition of k maps from $\mathcal{M}$. Let $I \subset J_z$ be a subarc such that $\mathcal{T}$ maps I diffeomorphically onto $J_{\mathcal{T}z}$. Then the conditional probability density $\psi_I = |\int_I \psi \, d\mathfrak{m}|^{-1} \psi 1_I$ satisfies*

$$\mathcal{L}_{\mathcal{T}}(\psi_I) \in \mathcal{D}_L,$$

the (possible) jump being located at $\mathcal{T}z$. If $k \geq N(L)$, then $\mathcal{L}_{\mathcal{T}}(\psi_I) \in \mathcal{D}_{L_*}$.

Proof. Let $L > 0$ and assume $\psi \in \mathcal{D}_L$, with z as in the definition above. Let $\mathcal{T}$ be a composition of k maps from $\mathcal{M}$, for some $k \geq 1$. We can subdivide J_z into intervals $I_1, \ldots, I_j$ such that $\tau_m = \mathcal{T}|I_m : I_m \to J_{\mathcal{T}z}$ is smooth and bijective for $1 \leq m \leq j$. Then

$$\mathcal{L}_{\mathcal{T}} \psi = \sum_{m=1}^{j} \mathcal{L}_{\tau_m} \psi,$$

where $\mathcal{L}_{\tau_m} \psi(x) = (\psi / \mathcal{T}')(\tau_m^{-1} x)$. Moreover, using (10), for some $K < \infty$,

$$\begin{aligned}
&|\log \mathcal{L}_{\tau_m} \psi(x) - \log \mathcal{L}_{\tau_m} \psi(y)| \\
&\quad \leq |\log \psi(\tau_m^{-1} x) - \log \psi(\tau_m^{-1} y)| + |\log \mathcal{T}'(\tau_m^{-1} x) - \log \mathcal{T}'(\tau_m^{-1} y)| \\
&\quad \leq L|\tau_m^{-1} x - \tau_m^{-1} y|_z + \|\mathcal{T}''/(\mathcal{T}')^2\|_\infty |x-y|_{\mathcal{T}z} \\
&\quad \leq (\lambda^{-k} L + K)|x-y|_{\mathcal{T}z} \equiv r.
\end{aligned}$$

Hence $e^{-r} \mathcal{L}_{\tau_m} \psi(x) \leq \mathcal{L}_{\tau_m} \psi(y) \leq e^r \mathcal{L}_{\tau_m} \psi(x)$. Summing over m yields

$$\mathcal{L}_{\mathcal{T}} \psi \in \mathcal{D}_{\lambda^{-k} L + K}, \quad \psi \in \mathcal{D}_L.$$

Fix $L_* = (1 - \lambda^{-1})^{-1} K > K$. Then parts (i) and (ii) of the lemma hold. As for part (iii), $\mathcal{L}_{\mathcal{T}}(\psi_I) = \mathcal{L}_{\tau_m}(\psi_I)$, for $I_m = I$. Since

$$|\log \mathcal{L}_{\mathcal{T}} \psi_I(x) - \log \mathcal{L}_{\mathcal{T}} \psi_I(y)| = |\log \mathcal{L}_{\tau_m} \psi(x) - \log \mathcal{L}_{\tau_m} \psi(y)| \leq r,$$

and since $\mathcal{L}_{\mathcal{T}} \psi_I$ is a probability density, we obtain part (iii). □

LEMMA 5.3. *There exists $C > 0$ such that*

$$\|\mathcal{L}_k \cdots \mathcal{L}_1 1\|_\infty \leq C, \tag{12}$$

$$\|\mathcal{L}_k \cdots \mathcal{L}_1\|_{C^0 \to C^0} \leq C, \tag{13}$$

$$\|\mathcal{L}_k \cdots \mathcal{L}_1\|_{\mathrm{Lip} \to \mathrm{Lip}} \leq C, \tag{14}$$

and

$$\|\mathcal{L}_k \cdots \mathcal{L}_1 h\|_{\mathrm{Lip}} \leq C \lambda^{-k} \|h\|_{\mathrm{Lip}} + C \|h\|_{C^0}. \tag{15}$$

https://doi.org/10.1017/etds.2016.9 Published online by Cambridge University Press

Proof. Lemma 5.2(i) implies $\mathcal{L}_k \cdots \mathcal{L}_1 1 \in \mathcal{D}_{L_*}$, which implies (12), from which (13) follows immediately. Now let h be Lipschitz continuous. Let $\mathcal{T} = T_k \circ \cdots \circ T_1$. Let $x_{i,-k}$ and $y_{i,-k}$ denote the preimages of x and y under the same (ith) branch of the inverse of $\mathcal{T}$. (The inverse branches are defined relative to the shorter arc connecting x and y.) Recalling (11), we deduce that

$$\begin{aligned}
|\mathcal{L}_{\mathcal{T}} h(x) - \mathcal{L}_{\mathcal{T}} h(y)| &\le \sum_i \left| \frac{h(x_{i,-k})}{\mathcal{T}'(x_{i,-k})} - \frac{h(y_{i,-k})}{\mathcal{T}'(y_{i,-k})} \right| \\
&\le \sum_i \left| \frac{h(x_{i,-k}) - h(y_{i,-k})}{\mathcal{T}'(x_{i,-k})} + \frac{h(y_{i,-k})}{\mathcal{T}'(y_{i,-k})} \left[\frac{\mathcal{T}'(y_{i,-k})}{\mathcal{T}'(x_{i,-k})} - 1 \right] \right| \\
&\le \lambda^{-k} \mathrm{Lip}(h) d(x, y) \sum_i \frac{1}{\mathcal{T}'(x_{i,-k})} + C d(x, y) \sum_i \frac{|h|(y_{i,-k})}{\mathcal{T}'(y_{i,-k})} \\
&\le (\lambda^{-k} \mathrm{Lip}(h) \mathcal{L}_{\mathcal{T}} 1(x) + C \mathcal{L}_{\mathcal{T}} |h|(y))\, d(x, y) \\
&\le (\lambda^{-k} \mathrm{Lip}(h) + C \|h\|_\infty) \|\mathcal{L}_{\mathcal{T}} 1\|_\infty\, d(x, y).
\end{aligned}$$

Noting that $\|\mathcal{L}_{\mathcal{T}} 1\|_\infty$ is uniformly bounded by (12),

$$\mathrm{Lip}(\mathcal{L}_{\mathcal{T}} h) \le C(\lambda^{-k} \mathrm{Lip}(h) + \|h\|_\infty),$$

which implies (15). Clearly, (14) is a consequence of (15). Let us note that, if h is allowed to have a jump at z, that is, if h is only assumed to be Lipschitz continuous with constant $L > 0$, say, on J_z,

$$|\mathcal{L}_{\mathcal{T}} h(x) - \mathcal{L}_{\mathcal{T}} h(y)| \le C(\lambda^{-k} L + \|h\|_\infty) |x - y|_{\mathcal{T}_z}. \tag{16}$$

□

LEMMA 5.4. *Suppose a probability density ψ is Lipschitz continuous on J_z with constant L for some $z \in \mathbb{S}$. Let $\mathcal{T}$ be a composition of k maps from $\mathcal{M}$. Then $\mathcal{L}_{\mathcal{T}} \psi$ is Lipschitz continuous on $J_{\mathcal{T}_z}$ with constant $C(1 + L)$. Here $C > 0$ is a uniform constant.*

Proof. Observing that $\|\psi\|_\infty \le 1 + L$, the result is a consequence of (16). □

LEMMA 5.5. *There exists a constant $C > 0$ such that, for any $T_1, T_2 \in \mathcal{M}$,*

$$\|\mathcal{L}_1 - \mathcal{L}_2\|_{\mathrm{Lip} \to C^0} \le C d_{C^1}(T_1, T_2). \tag{17}$$

In particular,

$$\|\mathcal{L}_t - \mathcal{L}_s\|_{\mathrm{Lip} \to C^0} \le C |t - s|^\eta. \tag{18}$$

Let $T_i, \tilde{T}_i \in \mathcal{M}$, $1 \le i \le k$ with associated transfer operators $\mathcal{L}_i, \tilde{\mathcal{L}}_i$. Then

$$\|\mathcal{L}_k \cdots \mathcal{L}_1 - \tilde{\mathcal{L}}_k \cdots \tilde{\mathcal{L}}_1\|_{\mathrm{Lip} \to C^0} \le C \sum_{j=1}^k d_{C^1}(T_j, \tilde{T}_j) \tag{19}$$

holds uniformly.

For obtaining perturbation estimates of the type in Lemma 5.5, it is paramount that the transfer operators act from a space of rather regular functions (here Lip) to a space of *less* regular functions (here C^0) (see [**14**, **17**] for a general theory).

https://doi.org/10.1017/etds.2016.9 Published online by Cambridge University Press

Proof of Lemma 5.5. Let $x \in \mathbb{S}$ and let $B \subset \mathbb{S}$ be the neighborhood of x of radius $\frac{1}{4}$. Set $w_T = \int_{\mathbb{S}} T' \, d\mathrm{m} \in \mathbb{N}$. For each $T \in \mathcal{M}$, the w_T inverse branches of T on B are well defined. When $d_{C^1}(T_1, T_2) < \frac{1}{8}$, there is a canonical correspondence between the inverse branches $T_{1,i}^{-1}$ and $T_{2,i}^{-1}$, $1 \le i \le w_{T_1}$ such that $T_{1,i}^{-1}(B)$ and $T_{2,i}^{-1}(B)$ overlap (in particular, $w_{T_1} = w_{T_2}$). Let us denote $y_i = T_{1,i}^{-1}(x)$ and $\tilde{y}_i = T_{2,i}^{-1}(x)$. Then $T_2(\tilde{y}_i) = T_1(y_i)$ implies that

$$d(y_i, \tilde{y}_i) \le \lambda^{-1} d(T_2(y_i), T_2(\tilde{y}_i)) = \lambda^{-1} d(T_2(y_i), T_1(y_i)).$$

Hence, the identity

$$\begin{aligned}(\mathcal{L}_1 - \mathcal{L}_2)h(x) &= \sum_{i=1}^{w} \left(\frac{h(y_i)}{T_1'(y_i)} - \frac{h(\tilde{y}_i)}{T_2'(\tilde{y}_i)} \right) \\ &= \sum_{i=1}^{w} \left[\frac{h(y_i) - h(\tilde{y}_i)}{T_1'(y_i)} + h(\tilde{y}_i) \left(\frac{1}{T_1'(y_i)} - \frac{1}{T_2'(\tilde{y}_i)} \right) \right] \\ &= \sum_{i=1}^{w} \left[\frac{h(y_i) - h(\tilde{y}_i)}{T_1'(y_i)} + h(\tilde{y}_i) \left(\frac{T_1'(\tilde{y}_i) - T_1'(y_i)}{T_1'(y_i) T_1'(\tilde{y}_i)} + \frac{T_2'(\tilde{y}_i) - T_1'(\tilde{y}_i)}{T_2'(\tilde{y}_i) T_1'(\tilde{y}_i)} \right) \right]\end{aligned}$$

yields, via (13) to obtain the uniform C,

$$\|(\mathcal{L}_1 - \mathcal{L}_2)h\|_\infty \le C \|h\|_{\mathrm{Lip}} d_{C^1}(T_1, T_2),$$

provided $d_{C^1}(T_1, T_2) < \frac{1}{8}$. For the case $d_{C^1}(T_1, T_2) \ge \frac{1}{8}$, we can use (12), giving

$$\|(\mathcal{L}_1 - \mathcal{L}_2)h\|_\infty \le 2\|h\|_\infty \sup_{T \in \mathcal{M}} \|\mathcal{L}_T 1\|_\infty \le C \|h\|_\infty \tfrac{1}{8} \le C \|h\|_{\mathrm{Lip}} d_{C^1}(T_1, T_2).$$

The above bounds prove (17). By Hölder continuity of γ, (18) clearly holds. It now follows from the identity

$$\mathcal{L}_k \cdots \mathcal{L}_1 - \tilde{\mathcal{L}}_k \cdots \tilde{\mathcal{L}}_1 = \sum_{j=1}^{k} \mathcal{L}_k \cdots \mathcal{L}_{j+1} (\mathcal{L}_j - \tilde{\mathcal{L}}_j) \tilde{\mathcal{L}}_{j-1} \cdots \tilde{\mathcal{L}}_1$$

and the uniform bounds in (13) and (14) that also (19) is satisfied. □

LEMMA 5.6. *There exists $\vartheta \in (0, 1)$ and, given $L > 0$, a constant $C > 0$ such that*

$$\|\mathcal{L}_k \cdots \mathcal{L}_1 (\psi^1 - \psi^2)\|_{L^1} \le C \vartheta^k, \quad k \ge 0$$

for all $\psi^1, \psi^2 \in \mathcal{D}_L$.

Proof. Note that

$$\psi \ge e^{-L_*}, \quad \psi \in \mathcal{D}_{L_*}.$$

Moreover, there exists $L_1 > 0$ such that

$$R(\psi) \equiv \frac{\psi - (1/2)e^{-L_*}}{1 - (1/2)e^{-L_*}} \in \mathcal{D}_{L_1}, \quad \psi \in \mathcal{D}_{L_*}.$$

We can now construct an inductive coupling scheme for exponential convergence. Suppose $L > 0$ is given and $\psi_i \in \mathcal{D}_L$, $i = 1, 2$. Denoting $\psi_k^i = \mathcal{L}_k \cdots \mathcal{L}_1 \psi^i$, $\psi_{N(L)}^i \ge e^{-L_*}$, by Lemma 5.2. Therefore,

$$\|\psi_{N(L)}^1 - \psi_{N(L)}^2\|_{L^1} = (1 - \tfrac{1}{2}e^{-L_*}) \|R(\psi_{N(L)}^1) - R(\psi_{N(L)}^2)\|_{L^1}.$$

https://doi.org/10.1017/etds.2016.9 Published online by Cambridge University Press

Moreover, $R(\psi^i_{N(L)}) \in \mathcal{D}_{L_1}$, $i = 1, 2$. Next, we repeat the procedure, treating $R(\psi^i_{N(L)})$ as the initial densities, obtaining another factor of $1 - \frac{1}{2}e^{-L_*}$ after $N(L_1)$ steps, and so on. This yields

$$\|\psi^1_k - \psi^2_k\|_{L^1} \le 2(1 - \tfrac{1}{2}e^{-L_*})^{j+1}$$

for any $k \ge N(L) + jN(L_1)$ and $j \ge 0$. Since the *a priori* bound $\|\psi^1_k - \psi^2_k\|_{L^1} \le 2$ holds for $k < N(L)$, the result follows. □

COROLLARY 5.7. *Given $L > 0$ and $z \in \mathbb{S}$, let $g : \mathbb{S} \to \mathbb{R}$ be a Lipschitz continuous function on J_z with constant L and assume that $\mathfrak{m}(g) = 0$.*

$$\|\mathcal{L}_k \cdots \mathcal{L}_1 g\|_{L^1} \le C\vartheta^k, \quad k \ge 0.$$

Here ϑ is the same constant as in Lemma 5.6 and C only depends on L.

Proof. Since $\mathfrak{m}(g) = 0$, $g \ge -L$. Then the probability density

$$\psi = \frac{g}{1+L} + 1$$

is Lipschitz continuous on J_z with constant $L/(1+L)$ and $\psi \ge 1/(1+L)$. By Remark 4.4, $\psi \in \mathcal{D}_L$. By Lemma 5.6,

$$\|\mathcal{L}_k \cdots \mathcal{L}_1(\psi - 1)\|_{L^1} \le C\vartheta^k,$$

where C is determined by the value of L. We get

$$\|\mathcal{L}_k \cdots \mathcal{L}_1 g\|_{L^1} \le C(1+L)\vartheta^k,$$

as claimed. □

Denoting $\hat{f}_t = f - \mathfrak{m}(\hat{\rho}_t f)$, Corollary 5.7, above, implies that

$$\|\mathcal{L}^k_t(\hat{\rho}_t \hat{f}_t)\|_{L^1} \le C\vartheta^k \tag{20}$$

holds uniformly in t and k, with the constant depending on $\|f\|_{\mathrm{Lip}}$ only.

Corollary 5.7 is also instrumental in the following lemma.

LEMMA 5.8. *For any $\beta > 0$, there exists a constant $C > 0$ such that, for any $T_1, T_2 \in \mathcal{M}$,*

$$\|\hat{\rho}_1 - \hat{\rho}_2\|_{L^1} \le C d_{C^1}(T_1, T_2)^{1-\beta}, \tag{21}$$

where $\hat{\rho}_j$ is the SRB density for T_j.

For any $\eta' < \eta$, there exists a constant $C > 0$ such that

$$\|\hat{\rho}_t - \hat{\rho}_s\|_{L^1} \le C|t - s|^{\eta'}. \tag{22}$$

Proof. We could appeal to the perturbation theory developed in [**17**] (see especially Corollary 1 and Remark 5 there). However, since we need control of $\hat{\rho}_1 - \hat{\rho}_2$ in L^1 only, we provide an independent argument. Note that $\hat{\rho}_T = \mathcal{L}^k_T \hat{\rho}_T$ for all $T \in \mathcal{M}$ and $k \ge 1$. Since the C^0 norm dominates the L^1 norm, the bounds in (19) and (8), together with Corollary 5.7, show that

$$\|\hat{\rho}_1 - \hat{\rho}_2\|_{L^1} \le \|\mathcal{L}^k_1(\hat{\rho}_1 - \hat{\rho}_2)\|_{L^1} + \|(\mathcal{L}^k_1 - \mathcal{L}^k_2)\hat{\rho}_2\|_{L^1} \le C(\vartheta^k + k d_{C^1}(T_1, T_2))$$

holds uniformly for all T_1, T_2 and all k. Choosing $k = \lceil \log d_{C^1}(T_1, T_2)/\log \vartheta \rceil$ on the right-hand side yields (21), from which (22) follows. □

https://doi.org/10.1017/etds.2016.9 Published online by Cambridge University Press

LEMMA 5.9. *There exists $b > 0$ such that the following holds. Given $\eta' < \eta$ and a probability density ρ that is Lipschitz continuous on J_z with constant $L > 0$ for some $z \in \mathbb{S}$,*

$$\|\rho_{n,k} - \hat{\rho}_{n,k}\|_{L^1} \le Cn^{-\eta'}, \quad k \ge b \log n.$$

Furthermore,

$$\|\rho_{n,\lfloor nt \rfloor} - \hat{\rho}_t\|_{L^1} \le Cn^{-\eta'} \tag{23}$$

and, more generally,

$$\|\rho_{n,\lfloor nt \rfloor} - \hat{\rho}_s\|_{L^1} \le C(n^{-\eta'} + |t-s|^{\eta'}), \tag{24}$$

provided $t \ge bn^{-1} \log n$. The constant $C > 0$ is determined by η' and L.

Proof. Since $\mathcal{L}_{n,k}\hat{\rho}_{n,k} = \hat{\rho}_{n,k}$,

$$\rho_{n,k} - \hat{\rho}_{n,k} = \mathcal{L}_{n,k} \cdots \mathcal{L}_{n,k-K+1}(\rho_{n,k-K} - \hat{\rho}_{n,k}) + (\mathcal{L}_{n,k} \cdots \mathcal{L}_{n,k-K+1} - \mathcal{L}_{n,k}^K)\hat{\rho}_{n,k}$$

for $1 \le K < k$. In order to bound $\rho_{n,k} - \hat{\rho}_{n,k}$ in L^1, note that $\rho_{n,k-K} - \hat{\rho}_{n,k}$ has a Lipschitz bound by Lemma 5.4, so the exponential bound in Corollary 5.7 implies

$$\|\mathcal{L}_{n,k} \cdots \mathcal{L}_{n,k-K+1}(\rho_{n,k-K} - \hat{\rho}_{n,k})\|_{L^1} \le C\vartheta^K,$$

where C is determined by L. Because $\|\hat{\rho}_{n,k}\|_{\mathrm{Lip}}$ is uniformly bounded (see (8)) and because $\|\cdot\|_{L^1} \le \|\cdot\|_{C^0}$, (19) applied to the difference $\mathcal{L}_{n,k} \cdots \mathcal{L}_{n,k-K+1} - \mathcal{L}_{n,k}^K$ yields

$$\|(\mathcal{L}_{n,k} \cdots \mathcal{L}_{n,k-K+1} - \mathcal{L}_{n,k}^K)\hat{\rho}_{n,k}\|_{L^1} \le CK \max_{k-K+1 \le j \le k} d_{C^1}(T_{n,j}, T_{n,k}) \le CK^{1+\eta}n^{-\eta}.$$

The last inequality is obtained from (1) and the Hölder continuity of γ, since

$$\begin{aligned} d_{C^1}(T_{n,j}, T_{n,k}) &\le d_{C^1}(T_{n,nn^{-1}j}, \gamma_{n^{-1}j}) + d_{C^1}(\gamma_{n^{-1}j}, \gamma_{n^{-1}k}) + d_{C^1}(\gamma_{n^{-1}k}, T_{n,nn^{-1}k}) \\ &\le Cn^{-\eta} + C|n^{-1}j - n^{-1}k|^{\eta} + Cn^{-\eta}. \end{aligned}$$

Collecting,

$$\|\rho_{n,k} - \hat{\rho}_{n,k}\|_{L^1} \le C(\vartheta^K + K^{1+\eta}n^{-\eta})$$

for all n, K and $k > K$. Setting $K = \lceil -\eta \log n / \log \vartheta \rceil$, we see that all terms on the right-hand side above are bounded by $Cn^{-\eta'}$, where C is determined by η' and L. Recalling (1), Lemma 5.8, concerning the perturbation of SRB densities, yields

$$\|\rho_{n,\lfloor nt \rfloor} - \hat{\rho}_{n,\lfloor nt \rfloor} + \hat{\rho}_{n,\lfloor nt \rfloor} - \hat{\rho}_t + \hat{\rho}_t - \hat{\rho}_s\|_{L^1} \le C(n^{-\eta'} + |t-s|^{\eta'}),$$

as required. □

Recall that $f_{n,k} = f \circ T_{n,k} \circ \cdots \circ T_{n,1}$ (see (2)). Corollary 5.7 is key also in the proof of the following decorrelation result (see [**32**] for a generalization).

LEMMA 5.10. *There exists a constant $\vartheta \in (0,1)$ such that the following holds. Given a probability measure μ with a density that is Lipschitz continuous on J_z with constant $L > 0$ for some $z \in \mathbb{S}$, $k \ge 2$, Lipschitz continuous functions $f^{(1)}, \ldots, f^{(k)} : \mathbb{S} \to \mathbb{R}$ and numbers $0 \le t_1 < \cdots < t_k \le 1$,*

$$\begin{aligned} &|\mu(f_{n,\lfloor nt_1 \rfloor}^{(1)} \cdots f_{n,\lfloor nt_k \rfloor}^{(k)}) - \mu(f_{n,\lfloor nt_1 \rfloor}^{(1)} \cdots f_{n,\lfloor nt_m \rfloor}^{(m)})\, \mu(f_{n,\lfloor nt_{m+1} \rfloor}^{(m+1)} \cdots f_{n,\lfloor nt_k \rfloor}^{(k)})| \\ &\quad \le C\vartheta^{n(t_{m+1}-t_m)} \end{aligned}$$

for all $m \in \{1, \ldots, k-1\}$ and $n \ge 0$. The constant $C > 0$ is determined by k, $\|f\|_{\mathrm{Lip}}$ and L.

https://doi.org/10.1017/etds.2016.9 Published online by Cambridge University Press

Proof. Let ρ be the density of μ. Let $z \in \mathbb{S}$ be such that ρ is Lipschitz on J_z with constant L. For the sake of brevity, let us denote

$$F = f^{(1)}_{n,\lfloor nt_1 \rfloor} \cdots f^{(m)}_{n,\lfloor nt_m \rfloor}, \qquad \tilde{F} = \mathcal{L}_{n,\lfloor nt_m \rfloor} \cdots \mathcal{L}_{n,1}(\rho F)$$

and

$$G = \prod_{j=m+1}^{k} f^{(j)} \circ T_{n,\lfloor nt_j \rfloor} \circ \cdots \circ T_{n,\lfloor nt_{m+1} \rfloor + 1}.$$

CLAIM. *The Lipschitz constant of $\tilde{F} - \mathfrak{m}(\tilde{F})$ on J_z and the function G are uniformly bounded over all n and t_j, $1 \le j \le k$.*

Assuming the claim,

$$\begin{aligned}
\mu(f^{(1)}_{n,\lfloor nt_1 \rfloor} \cdots f^{(k)}_{n,\lfloor nt_k \rfloor}) &= \mathfrak{m}(\rho F \cdot G \circ T_{n,\lfloor nt_{m+1} \rfloor} \circ \cdots \circ T_{n,1}) \\
&= \mathfrak{m}(\tilde{F} \cdot G \circ T_{n,\lfloor nt_{m+1} \rfloor} \circ \cdots \circ T_{n,\lfloor nt_m \rfloor + 1}) \\
&= \mathfrak{m}(\tilde{F})\mathfrak{m}(G \circ T_{n,\lfloor nt_{m+1} \rfloor} \circ \cdots \circ T_{n,\lfloor nt_m \rfloor + 1}) + O(\vartheta^{n(t_{m+1}-t_m)}) \\
&= \mu(F)\mathfrak{m}(G \circ T_{n,\lfloor nt_{m+1} \rfloor} \circ \cdots \circ T_{n,\lfloor nt_m \rfloor + 1}) + O(\vartheta^{n(t_{m+1}-t_m)}) \\
&= \mu(F)\mu_{n,\lfloor nt_m \rfloor}(G \circ T_{n,\lfloor nt_{m+1} \rfloor} \circ \cdots \circ T_{n,\lfloor nt_m \rfloor + 1}) \\
&\quad + O(\vartheta^{n(t_{m+1}-t_m)}) \\
&= \mu(F)\mu(G \circ T_{n,\lfloor nt_{m+1} \rfloor} \circ \cdots \circ T_{n,1}) + O(\vartheta^{n(t_{m+1}-t_m)}).
\end{aligned}$$

The first, fourth and sixth lines use definitions only, and the second line uses duality. The third and fifth lines follow from Corollary 5.7; the claim that G is bounded is needed for both, while the uniform Lipschitz bound on $\tilde{F} - \mathfrak{m}(\tilde{F})$ is crucial for the third line. In the fifth line, we also used the fact that the density $\rho_{n,\lfloor nt_m \rfloor}$ of $\mu_{n,\lfloor nt_m \rfloor}$ has a uniform bound on its Lipschitz constant, by Lemma 5.4; therefore $\mathfrak{m}(|\mathcal{L}_{n,\lfloor nt_{m+1} \rfloor} \cdots \mathcal{L}_{n,\lfloor nt_m \rfloor + 1}(1 - \rho_{n,\lfloor nt_m \rfloor})|) \le C\vartheta^{n(t_{m+1}-t_m)}$, by Corollary 5.7 .

Proof of Claim. The boundedness of G is obvious. Write $\mathcal{T} = T_{n,\lfloor nt_m \rfloor} \circ \cdots \circ T_{n,1}$. Since ρ is a probability density, its Lipschitz property implies the upper bound $\rho \le 1 + L$. Thus,

$$\|\rho F\|_\infty \le (1 + L) \prod_{j=1}^{m} \|f^{(j)}\|_\infty.$$

Given $x, y \in J_z$, we write $x_{i,-n}, y_{i,-n}$ for the corresponding preimages under $\mathcal{T}$. (In other words, for each i there is an arc $J_{i,-n}$ containing $x_{i,-n}$ and $y_{i,-n}$ such that $\mathcal{T}(J_{i,-n}) = J_z$.) The trivial bounds $|\rho(x_{i,-n}) - \rho(y_{i,-n})| \le L|x - y|_z$ and $|f^{(j)}_{n,\lfloor nt_j \rfloor}(x_{i,-n}) - f^{(j)}_{n,\lfloor nt_j \rfloor}(x_{i,-n})| \le \mathrm{Lip}(f^{(j)})|x - y|_z$ yield

$$|(\rho F)(x_{i,-n}) - (\rho F)(y_{i,-n})| \le |x - y|_z(1 + 2L) \prod_{j=1}^{m} \|f^{(j)}\|_{\mathrm{Lip}}.$$

https://doi.org/10.1017/etds.2016.9 Published online by Cambridge University Press

As in the proof of (15),

$$\begin{aligned}|\tilde{F}(x)-\tilde{F}(y)| &\le \sum_i \left|\frac{(\rho F)(x_{i,-n})}{\mathcal{T}'(x_{i,-n})}-\frac{(\rho F)(y_{i,-n})}{\mathcal{T}'(y_{i,-n})}\right| \\ &\le \sum_i \left|\frac{(\rho F)(x_{i,-n})-(\rho F)(y_{i,-n})}{\mathcal{T}'(x_{i,-n})}+\frac{(\rho F)(y_{i,-n})}{\mathcal{T}'(y_{i,-n})}\left[\frac{\mathcal{T}'(y_{i,-n})}{\mathcal{T}'(x_{i,-n})}-1\right]\right| \\ &\le C|x-y|_z(1+2L)\prod_{j=1}^{m}\|f^{(j)}\|_{\mathrm{Lip}},\end{aligned}$$

which implies the claim. The proof of Lemma 5.10 is complete. □

6. *Preliminaries II: processes ζ_n, ξ_n and χ_n*

In this section, we study the stochastic processes of interest, assuming the initial measure μ has density $\rho \in \bigcup_{L>0}\mathcal{D}_L$. Without mentioning it separately each time, the observable $f : \mathbb{S}\to\mathbb{R}$ is assumed to be Lipschitz continuous in the rest of the paper. For convenience, let us now recall the notation $f_{n,k}=f\circ T_{n,k}\circ\cdots\circ T_{n,1}$ and $\bar{f}_{n,k}=f_{n,k}-\mu(f_{n,k})$, introduced in (2) and (9), respectively.

Convention. Recall the definition of χ_n in (5). Given the initial measure μ, we will often consider the special centering $c_n(t)=\mu(\zeta_n(\cdot\,,t))$ which yields $\mu(\chi_n(\cdot\,,t))=0$ for all t. In this case, to emphasize that the initial measure has been chosen and appears in the definition of the process explicitly, *we use the symbol ξ_n instead of χ_n*: that is, we define the functions $\xi_n : \mathbb{S}\times[0,1]\to\mathbb{R}$ as

$$\xi_n(x,t)=n^{1/2}\zeta_n(x,t)-n^{1/2}\mu(\zeta_n(\cdot\,,t)). \tag{25}$$

In other words, given another centering sequence $(c_n)_{n\ge1}$, we have the relation

$$\chi_n(x,t)=\xi_n(x,t)+n^{1/2}\int_0^t \mu(f_{n,\lfloor ns\rfloor})\,ds-n^{1/2}c_n(x,t).$$

In practice, it will be convenient to express the processes ζ_n and ξ_n using integral notation: that is,

$$\zeta_n(x,t)=\int_0^t f_{n,\lfloor ns\rfloor}(x)\,ds \tag{26}$$

and

$$\xi_n(x,t)=n^{1/2}\int_0^t \bar{f}_{n,\lfloor ns\rfloor}(x)\,ds \tag{27}$$

from (3) and (25), respectively. From here on, we will routinely drop the x-dependence from the notation, writing just $\zeta_n(t)$ instead of $\zeta_n(x,t)$, and so on. Of course,

$$\xi_n(t_2)-\xi_n(t_1)=n^{1/2}\int_{t_1}^{t_2}\bar{f}_{n,\lfloor ns\rfloor}\,ds. \tag{28}$$

Let us record a useful bound: given any $\eta'\in(\frac{1}{2},\eta)$, $\mu_{n,\lfloor ns\rfloor}(f)=\mu(f_{n,\lfloor ns\rfloor})$ and (23) imply

$$\mu(f_{n,\lfloor ns\rfloor})-\hat{\mu}_s(f)=O(n^{-\eta'}),\quad s\ge bn^{-1}\log n. \tag{29}$$

https://doi.org/10.1017/etds.2016.9 Published online by Cambridge University Press

6.1. *The variance $\hat{\sigma}_t^2(f)$.*

LEMMA 6.1. *Recall that $\hat{f}_t = f - \hat{\mu}_t(f)$. The limit variance in (7) can be expressed in terms of the series*

$$\hat{\sigma}_t^2(f) = \hat{\mu}_t[\hat{f}_t^2] + 2\sum_{k=1}^{\infty} \hat{\mu}_t[\hat{f}_t\, \hat{f}_t \circ \gamma_t^k] = \hat{\mu}_t[\hat{f}_t^2] + 2\sum_{k=1}^{\infty} \mathfrak{m}[\hat{f}_t \mathcal{L}_t^k(\hat{\rho}_t\, \hat{f}_t)]. \tag{30}$$

The map $[0, 1] \to \mathbb{R}_+ : t \mapsto \hat{\sigma}_t^2(f)$ is (uniformly) continuous.

Proof. Notice that the two series are equal term by term. By (20),

$$\sup_{0\le s\le 1} |\mathfrak{m}[\hat{f}_s \mathcal{L}_t^k(\hat{\rho}_s\, \hat{f}_s)]| \le C\vartheta^k,$$

so the series converge absolutely at an exponential rate. On the other hand, a direct manipulation of (7) shows that

$$\hat{\sigma}_t^2(f) = \hat{\mu}_t(\hat{f}_t^2) + 2\lim_{m\to\infty} m^{-1} \sum_{k=1}^{m-1} (m-k)\hat{\mu}_t(\hat{f}_t\, \hat{f}_t \circ \gamma_t^k),$$

which suffices to prove (30).

Defining the truncated sum

$$V_{K,t} = \hat{\mu}_t[\hat{f}_t^2] + 2\sum_{k=1}^{K} \mathfrak{m}[\hat{f}_t \mathcal{L}_t^k(\hat{\rho}_t\, \hat{f}_t)],$$

we have

$$|\hat{\sigma}_t^2(f) - \hat{\sigma}_s^2(f)| \le |V_{K,t} - V_{K,s}| + C\vartheta^K$$

for all $s \in [0, 1]$ and all $K > 0$. Given $\varepsilon > 0$, we fix K so large that $C\vartheta^K < \varepsilon/2$. Note that $V_{K,t} = [\mathfrak{m}(\hat{\rho}_t f^2) - \mathfrak{m}(\hat{\rho}_t f)^2] + 2\sum_{k=1}^{K}[\mathfrak{m}(f\mathcal{L}_t^k(\hat{\rho}_t f)) - \mathfrak{m}(\hat{\rho}_t f)^2]$. Moreover,

$$\begin{aligned}
&|\mathfrak{m}(f\mathcal{L}_t^k(\hat{\rho}_t f)) - \mathfrak{m}(f\mathcal{L}_s^k(\hat{\rho}_s f))| \\
&\quad \le \|f\|_\infty \|\mathcal{L}_t^k(\hat{\rho}_t f) - \mathcal{L}_s^k(\hat{\rho}_s f)\|_{L^1} \\
&\quad \le \|f\|_\infty \|(\mathcal{L}_t^k - \mathcal{L}_s^k)(\hat{\rho}_t f)\|_{L^1} + \|f\|_\infty \|\mathcal{L}_s^k(\hat{\rho}_t f - \hat{\rho}_s f)\|_{L^1} \\
&\quad \le \|f\|_\infty \|\mathcal{L}_t^k - \mathcal{L}_s^k\|_{\mathrm{Lip}\to C^0} \|\hat{\rho}_t f\|_{\mathrm{Lip}} + \|f\|_\infty \|(\hat{\rho}_t - \hat{\rho}_s) f\|_{L^1} \\
&\quad \le \|f\|_{\mathrm{Lip}}^2 (\|\mathcal{L}_t^k - \mathcal{L}_s^k\|_{\mathrm{Lip}\to C^0} \|\hat{\rho}_t\|_{\mathrm{Lip}} + \|\hat{\rho}_t - \hat{\rho}_s\|_{L^1}).
\end{aligned}$$

By (8) of Remark 4.1, (19) of Lemma 5.5 and (22) of Lemma 5.8, we see that $|V_{K,t} - V_{K,s}| < \varepsilon/2$ and $|\hat{\sigma}_t^2(f) - \hat{\sigma}_s^2(f)| < \varepsilon$ for any s such that $|t - s|$ is sufficiently small. This proves continuity. □

6.2. *The second moment $\mu[[\xi_n(t + h) - \xi_n(t)]^2]$.*

LEMMA 6.2. *Let $L > 0$ and $\rho \in \mathcal{D}_L$. The quantity $\mu[[\xi_n(t) - \xi_n(s)]^2]$ is uniformly bounded over all $t, s \in [0, 1]$ and $n \ge 1$. Moreover,*

$$\mu[[\xi_n(t + h) - \xi_n(t)]^2] = \int_t^{t+h} \hat{\sigma}_s^2(f)\, ds + o(n^{-1/2}) + h\, o(1), \tag{31}$$

https://doi.org/10.1017/etds.2016.9 Published online by Cambridge University Press

as $n \to \infty$, whenever $0 \le t \le t+h \le 1$. The error terms are uniform in t and h†. Finally,

$$\int_t^{t+h} \hat{\sigma}_s^2(f)\, ds = h\hat{\sigma}_t^2(f) + o(h), \tag{32}$$

as $h \to 0$. The error term is uniform in t. All of the bounds depend on ρ through L only.

Proof. Given $L > 0$, the maximum of ρ is bounded by a constant determined by L. Because $\mu(\bar{f}_{n,k}) = 0$,

$$\mu[[\xi_n(t+h) - \xi_n(t)]^2] = n \int_t^{t+h} \int_t^{t+h} \mu(\bar{f}_{n,\lfloor ns \rfloor} \bar{f}_{n,\lfloor nr \rfloor})\, dr\, ds, \tag{33}$$

as $n \to \infty$. By Lemma 5.10,

$$|\mu(\bar{f}_{n,\lfloor ns \rfloor} \bar{f}_{n,\lfloor nr \rfloor})| \le C\vartheta^{n|r-s|}. \tag{34}$$

The boundedness claim of the lemma now follows from elementary integration: that is,

$$\begin{aligned} n \int_t^{t+h} \int_t^{t+h} \vartheta^{n|r-s|}\, dr\, ds &= 2n \int_0^h \int_s^h \vartheta^{n(r-s)}\, dr\, ds \\ &= \frac{2}{\log \vartheta} \int_0^h (\vartheta^{n(h-s)} - 1)\, ds \\ &\le \frac{2h}{|\log \vartheta|}. \end{aligned}$$

As to the second claim, let $\kappa \in (0, \frac{1}{4})$ be small enough that $2\kappa < \eta'(1-\kappa)$ and set $a_n = n^{-1+\kappa}$. We shall now show that, in the limit $n \to \infty$, the sole contribution to the double integral $n \int_t^{t+h} \int_t^{t+h} \mu(\bar{f}_{n,\lfloor ns \rfloor} \bar{f}_{n,\lfloor nr \rfloor})\, dr\, ds$ comes from the parallelepiped $P_n = \{(s, r) \in [t, t+h]^2 : t + 2a_n \le s \le t + h - a_n \text{ and } |r - s| \le a_n\}$ about the diagonal. To that end, let

$$Q_n = \{(s, r) \in [t, t+h]^2 : |r - s| \le a_n \text{ and either } s < t + 2a_n \text{ or } s > t + h - a_n\}$$

and let $R_n = \{(s, r) \in [t, t+h]^2 : |r - s| > a_n\}$, so that $[t, t+h]^2 = P_n \cup Q_n \cup R_n$. Because the area of Q_n is $O(a_n^2)$,

$$n \iint_{Q_n} \mu(\bar{f}_{n,\lfloor ns \rfloor} \bar{f}_{n,\lfloor nr \rfloor})\, dr\, ds = O(na_n^2) = O(n^{-1+2\kappa}).$$

On the other hand, (34) yields

$$n \iint_{R_n} \mu(\bar{f}_{n,\lfloor ns \rfloor} \bar{f}_{n,\lfloor nr \rfloor})\, dr\, ds = O(n\vartheta^{na_n}) = o(n^{-1}).$$

Thus, only the contribution of P_n is significant and

$$\begin{aligned} &n \int_t^{t+h} \int_t^{t+h} \mu(\bar{f}_{n,\lfloor ns \rfloor} \bar{f}_{n,\lfloor nr \rfloor})\, dr\, ds \\ &\quad = n \int_{t+2a_n}^{t+h-a_n} \int_{s-a_n}^{s+a_n} \mu(\bar{f}_{n,\lfloor ns \rfloor} \bar{f}_{n,\lfloor nr \rfloor})\, dr\, ds + O(na_n^2). \end{aligned} \tag{35}$$

† We remark that the exponent in $o(n^{-1/2})$ is not optimal, but a choice of convenience.

https://doi.org/10.1017/etds.2016.9 Published online by Cambridge University Press

Note that, on P_n, $s - a_n \geq t + a_n \geq a_n \geq bn^{-1} \log n$ for all but a finite number of n. (This motivates the odd $2a_n$ in the definition of P_n.) By (24),

$$\sup_{r \in (s-a_n, s+a_n)} \|\rho_{n,\lfloor nr \rfloor} - \hat{\rho}_s\|_{L^1} = O(n^{-\eta'} + a_n^{\eta'}) = O(a_n^{\eta'}), \tag{36}$$

so

$$\sup_{r \in (s-a_n, s+a_n)} |\mu(f_{n,\lfloor nr \rfloor}) - \hat{\mu}_s(f)| = O(a_n^{\eta'}).$$

This implies that

$$n \int_{s-a_n}^{s+a_n} \mu(\bar{f}_{n,\lfloor ns \rfloor} \bar{f}_{n,\lfloor nr \rfloor})\, dr = n \int_{s-a_n}^{s+a_n} \mu(f_{n,\lfloor ns \rfloor} f_{n,\lfloor nr \rfloor}) - \hat{\mu}_s(f)^2\, dr + O(n a_n^{1+\eta'}). \tag{37}$$

We split the domain of integration $[s - a_n, s + a_n]$ on the right-hand side into two halves. Setting $b_n = (1/n)(1 - \{ns\})$ (using (36) to pass to the third line),

$$\begin{aligned}
&n \int_s^{s+a_n} \mu(f_{n,\lfloor ns \rfloor} f_{n,\lfloor nr \rfloor})\, dr \\
&\quad = n \int_0^{a_n} \mu(f_{n,\lfloor ns \rfloor} f_{n,\lfloor n(s+r) \rfloor})\, dr \\
&\quad = b_n n \mu_{n,\lfloor ns \rfloor}(f^2) + n \int_{b_n}^{a_n} \mu_{n,\lfloor ns \rfloor}(ff \circ T_{n,\lfloor n(s+r) \rfloor} \circ \cdots \circ T_{n,\lfloor ns \rfloor + 1})\, dr \\
&\quad = b_n n \hat{\mu}_s(f^2) + n \int_{b_n}^{a_n} \hat{\mu}_s(ff \circ T_{n,\lfloor n(s+r) \rfloor} \circ \cdots \circ T_{n,\lfloor ns \rfloor + 1})\, dr + O(a_n^{\eta'} + n a_n^{1+\eta'}) \\
&\quad = n \int_0^{b_n} \mathfrak{m}(f \hat{\rho}_s f)\, dr + n \int_{b_n}^{a_n} \mathfrak{m}(f \mathcal{L}_{n,\lfloor n(s+r) \rfloor} \cdots \mathcal{L}_{n,\lfloor ns \rfloor + 1}(\hat{\rho}_s f))\, dr + O(n a_n^{1+\eta'}).
\end{aligned}$$

We can replace $\mathcal{L}_{n,\lfloor n(s+r) \rfloor} \cdots \mathcal{L}_{n,\lfloor ns \rfloor + 1}$ by $\mathcal{L}_s^{\lfloor n(s+r) \rfloor - \lfloor ns \rfloor}$ since, recalling (19) and (8),

$$\begin{aligned}
&\|\mathcal{L}_{n,\lfloor n(s+r) \rfloor} \cdots \mathcal{L}_{n,\lfloor ns \rfloor + 1}(\hat{\rho}_s f) - \mathcal{L}_s^{\lfloor n(s+r) \rfloor - \lfloor ns \rfloor}(\hat{\rho}_s f)\|_{L^1} \\
&\quad \leq \|\mathcal{L}_{n,\lfloor n(s+r) \rfloor} \cdots \mathcal{L}_{n,\lfloor ns \rfloor + 1} - \mathcal{L}_s^{\lfloor n(s+r) \rfloor - \lfloor ns \rfloor}\|_{\mathrm{Lip} \to C^0} \|\hat{\rho}_s f\|_{\mathrm{Lip}} \\
&\quad \leq C n a_n^{1+\eta} = O(n a_n^{1+\eta'})
\end{aligned}$$

uniformly for $r \in [0, a_n]$. Hence,

$$n \int_s^{s+a_n} \mu(f_{n,\lfloor ns \rfloor} f_{n,\lfloor nr \rfloor})\, dr = n \int_0^{a_n} \mathfrak{m}(f \mathcal{L}_s^{\lfloor n(s+r) \rfloor - \lfloor ns \rfloor}(\hat{\rho}_s f))\, dr + O(n^2 a_n^{2+\eta'}).$$

A similar computation, which we leave to the reader, yields

$$n \int_{s-a_n}^{s} \mu(f_{n,\lfloor ns \rfloor} f_{n,\lfloor nr \rfloor})\, dr = n \int_{-a_n}^{0} \mathfrak{m}(f \mathcal{L}_s^{\lfloor ns \rfloor - \lfloor n(s+r) \rfloor}(\hat{\rho}_s f))\, dr + O(n^2 a_n^{2+\eta'}).$$

By (37), we have thus shown that

$$\begin{aligned}
&n \int_{s-a_n}^{s+a_n} \mu(\bar{f}_{n,\lfloor ns \rfloor} \bar{f}_{n,\lfloor nr \rfloor})\, dr \\
&\quad = n \int_{-a_n}^{a_n} \mathfrak{m}(\hat{f}_s \mathcal{L}_s^{|\lfloor n(s+r) \rfloor - \lfloor ns \rfloor|}(\hat{\rho}_s \hat{f}_s))\, dr + O(n^2 a_n^{2+\eta'}) \\
&\quad = n \int_{-\infty}^{\infty} \mathfrak{m}(\hat{f}_s \mathcal{L}_s^{|\lfloor n(s+r) \rfloor - \lfloor ns \rfloor|}(\hat{\rho}_s \hat{f}_s))\, dr + O(\vartheta^{n a_n} + n^2 a_n^{2+\eta'}) \\
&\quad = \hat{\sigma}_s^2(f) + O(n^2 a_n^{2+\eta'}).
\end{aligned}$$

https://doi.org/10.1017/etds.2016.9 Published online by Cambridge University Press

The second from last line follows from $|\mathfrak{m}(\hat{f}_s\mathcal{L}_s^k(\hat{\rho}_s\hat{f}_s))| \le C\|\mathcal{L}_s^k(\hat{\rho}_s\hat{f}_s)\|_{L^1} \le C\vartheta^k$ (see (20)) and the last one from (30). Recalling (35) and (33), we obtain

$$\begin{aligned}\mu[[\xi_n(t+h)-\xi_n(t)]^2] &= \int_{t+2a_n}^{t+h-a_n} \hat{\sigma}_s^2(f)\,ds + O(na_n^2 + hn^2a_n^{2+\eta'}) \\ &= \int_t^{t+h} \hat{\sigma}_s^2(f)\,ds + O(a_n + na_n^2 + hn^2a_n^{2+\eta'}) \\ &= \int_t^{t+h} \hat{\sigma}_s^2(f)\,ds + O(n^{-1+2\kappa} + hn^{2\kappa-\eta'(1-\kappa)}),\end{aligned}$$

which implies (31), by choice of κ.

Finally, the function $s \mapsto \hat{\sigma}_s^2(f)$ is uniformly continuous, by Lemma 6.1. Hence, it has an increasing modulus of continuity $w : [0, 1] \to \mathbb{R}_+$ such that $|\hat{\sigma}_t^2(f) - \hat{\sigma}_s^2(f)| \le w(|t-s|)$ holds for all $s, t \in [0, 1]$ and $\lim_{\delta\to 0} w(\delta) = 0$. Therefore, $|\int_t^{t+h} \hat{\sigma}_s^2(f)\,ds - h\hat{\sigma}_t^2(f)| \le h\,w(h) = o(h)$ as $h \to 0$, uniformly in t. □

6.3. *Decorrelation at the process level.* Next, we introduce useful partitions of $\mathbb{S}$ having the property that $x \mapsto \xi_n(x, t)$ (and $x \mapsto \zeta_n(x, t)$) is nearly constant on each partition element. To that end, fix $z \in \mathbb{S}$. (The role of z is to define branches and to keep track of a possible discontinuity of a probability density; see §§4 and 5.) For any integer $n \ge 1$ and real number $t \in (0, 1)$, there are $N_{n,t}$ points y_j, $j = 1, \dots, N_{n,t}$ with the same image $z_{n,t}$ as z under $T_{n,\lfloor nt\rfloor} \circ \cdots \circ T_{n,1}$, where $N_{n,t}$ is the degree of the composition. The points $\{y_j\}_{j=1}^{N_{n,t}}$ define an induced partition $\mathcal{P}_{z,n,t} = \{I_{z,n,t,j}\}_{j=1}^{N_{n,t}}$ of $\mathbb{S}\backslash\{y_j : j = 1, \dots, N_{n,t}\}$ into subarcs $I_{z,n,t,j}$ with the property that the restriction of $T_{n,\lfloor nt\rfloor} \circ \cdots \circ T_{n,1}$ to $I_{z,n,t,j}$ is one-to-one and onto $J_{z_{n,t}} = \mathbb{S}\backslash\{z_{n,t}\}$ for all j. It follows, from the uniform expansion property, that

$$|T_{n,\lfloor ns\rfloor} \circ \cdots \circ T_{n,1}(I_{z,n,t,j})| \le C\lambda^{n(s-t)}$$

for all $s \le t$. Since $f \in \mathrm{Lip}$, this yields the uniform bound

$$|\xi_n(x, s) - \xi_n(y, s)| \le Cn^{-1/2} \tag{38}$$

for all $x, y \in I_{z,n,t,j}$ and j, for all $s \le t$. Given bounded and Lipschitz continuous functions $B_1, \dots, B_m$ on $\mathbb{R}$, real numbers $0 \le t_1 < \cdots < t_m \le t$ and a probability measure μ with density $\rho > 0$, integrating the previous bound with respect to $d\mu(y)$ yields the existence of C—determined by $\prod_{1\le k\le m} \|B_k\|_{\mathrm{Lip}}$—such that

$$|B_1(\xi_n(x, t_1)) \dots B_m(\xi_n(x, t_m)) - \mu_{z,n,t,j}[B_1(\xi_n(t_1)) \dots B_m(\xi_n(t_m))]| \le Cn^{-1/2} \tag{39}$$

for all $x \in I_{z,n,t,j}$ and j. Here $\mu_{z,n,t,j}$ denotes the conditional measure $(1/\mu(I_{z,n,t,j}))\,\mu[1_{I_{z,n,t,j}} \cdot\,]$.

LEMMA 6.3. *Suppose* $\rho \in \bigcup_{L>0} \mathcal{D}_L$.

(i) *If* $A \in C^\infty(\mathbb{R})$, *then*

$$\mu[A(\zeta_n(s))[\zeta_n(t) - \zeta_n(s)]] - \mu[A(\zeta_n(s))]\mu[\zeta_n(t) - \zeta_n(s)] = o(1)$$

as $n \to \infty$, *whenever* $0 \le s \le t \le 1$.

https://doi.org/10.1017/etds.2016.9 Published online by Cambridge University Press

(ii) *If* $A \in C_c^\infty(\mathbb{R})$ *and* $q \in \{1, 2\}$, *then*

$$\mu[A(\xi_n(s))[\xi_n(t) - \xi_n(s)]^q] - \mu[A(\xi_n(s))]\mu[[\xi_n(t) - \xi_n(s)]^q] = o(1)$$

as $n \to \infty$, *whenever* $0 \le s \le t \le 1$.

Proof. Note that, since ζ_n is uniformly bounded, it suffices to assume that $A \in C_c^\infty(\mathbb{R})$ in both parts of the lemma. We only prove part (ii), concerning ξ_n, and leave the easier, but similar, part (i), concerning ζ_n, to the reader.

Let $z \in \mathbb{S}$ be a point such that $\log \rho$ is Lipschitz continuous on J_z with constant $L > 0$, and consider the induced partition $\mathcal{P}_{z,n,s} = \{I_{z,n,s,j}\}_{j=1}^{N_{n,s}}$ relative to this point. Since $\mu[[\xi_n(t) - \xi_n(s)]^q]$ is uniformly bounded, by Lemma 6.2, a special case of (39) yields

$$\begin{aligned}
&\mu[A(\xi_n(s))[\xi_n(t) - \xi_n(s)]^q] \\
&\quad = \sum_j \mu[1_{I_{z,n,s,j}} A(\xi_n(s))[\xi_n(t) - \xi_n(s)]^q] \\
&\quad = \sum_j \mu_{z,n,s,j}[A(\xi_n(s))]\, \mu[1_{I_{z,n,s,j}}[\xi_n(t) - \xi_n(s)]^q] + O(n^{-1/2}) \\
&\quad = \sum_j \mu[1_{I_{z,n,s,j}} A(\xi_n(s))]\, \mu_{z,n,s,j}[[\xi_n(t) - \xi_n(s)]^q] + O(n^{-1/2}).
\end{aligned}$$

Let us denote the density of the conditional measure $\mu_{z,n,s,j}$ by $\rho_{z,n,s,j}$. To finish the proof, it is enough to show that

$$\max_j |\mu_{z,n,s,j}[[\xi_n(t) - \xi_n(s)]^q] - \mu[[\xi_n(t) - \xi_n(s)]^q]| = o(1)$$

as $n \to \infty$. Fixing $p > \frac{1}{2}$, arbitrarily, it follows, from (28), that

$$\xi_n(t) - \xi_n(s) = n^{1/2} \int_{s+n^{-p}}^t \bar{f}_{n,\lfloor nr \rfloor}\, dr + o(1)$$

holds for any t and s. In other words, the lower limit of integration can be slightly increased, and it is sufficient to prove

$$\max_j \left| \mu_{z,n,s,j}\left[\left[n^{1/2} \int_{s+n^{-p}}^t \bar{f}_{n,\lfloor nr \rfloor}\, dr\right]^q\right] - \mu\left[\left[n^{1/2} \int_{s+n^{-p}}^t \bar{f}_{n,\lfloor nr \rfloor}\, dr\right]^q\right] \right| = o(1).$$

The increment n^{-p} facilitates the use of L^1 convergence of the pushforwards of $\rho_{z,n,s,j}$ and ρ. Indeed, since $\bar{f}_{n,\lfloor nr \rfloor}$ is uniformly bounded, it suffices to show that

$$\max_j \|\mathcal{L}_{n,\lfloor n(s+n^{-p}) \rfloor} \cdots \mathcal{L}_{n,1}(\rho_{z,n,s,j} - \rho)\|_{L^1} = o(n^{-q/2}).$$

Without loss of generality, we may assume that $L \ge L_*$. By virtue of Lemma 5.2, both of the densities $\tilde{\rho}_{z,n,s,j} = \mathcal{L}_{n,\lfloor ns \rfloor} \cdots \mathcal{L}_{n,1} \rho_{z,n,s,j}$ and $\rho_{n,\lfloor ns \rfloor} = \mathcal{L}_{n,\lfloor ns \rfloor} \cdots \mathcal{L}_{n,1} \rho$ are then in $\mathcal{D}_L$. Lemma 5.6 now yields

$$\max_j \|\mathcal{L}_{n,\lfloor n(s+n^{-p}) \rfloor} \cdots \mathcal{L}_{n,\lfloor ns \rfloor + 1}(\tilde{\rho}_{z,n,s,j} - \rho_{n,\lfloor ns \rfloor})\|_{L^1} \le C\vartheta^{n^{1-p}},$$

where C depends on L. The proof is complete. □

https://doi.org/10.1017/etds.2016.9 Published online by Cambridge University Press

7. *Proof of Theorem 3.1*

Recall the integral expression of ζ_n in (26). Given an initial probability measure μ on $\mathbb{S}$, each ζ_n is a random element of $C^0([0,1],\mathbb{R})$ with distribution $\mathbf{P}_n^\mu$. The corresponding expectation will be denoted by $\mathbf{E}_n^\mu$.

Recall that a tight sequence of measures has limit points in the topology of weak convergence [**5**].

LEMMA 7.1. *Let the measure μ be arbitrary. The sequence of measures $(\mathbf{P}_n^\mu)_{n\geq 1}$ is tight.*

Proof. Note that, for $t_1, t_2 \in [0,1]$,

$$\zeta_n(t_2) - \zeta_n(t_1) = \int_{t_1}^{t_2} f_{n,\lfloor ns\rfloor}\, ds.$$

Accordingly, $|\zeta_n(t_2) - \zeta_n(t_1)| \leq (t_2 - t_1)\|f\|_\infty$. In other words, the sequence $(\zeta_n)_{n\geq 1}$ of functions is uniformly Lipschitz and bounded. This suffices for tightness on the classical Wiener space $C^0([0,1],\mathbb{R})$. □

Next, we are going to show that the sequence actually has a unique limit, which we are going to identify. The following Dynkin formula will turn out to be useful in this regard. In order to formulate it properly, let us introduce the evaluation functionals $\pi_t : C^0([0,1],\mathbb{R}) \to \mathbb{R}$, $t \in [0,1]$, defined by

$$\pi_t(\omega) = \omega(t).$$

LEMMA 7.2. *Let $\rho \in \bigcup_{L>0} \mathcal{D}_L$. Suppose that $\mathbf{P}$ is the weak limit of a subsequence $(\mathbf{P}_{n_k}^\mu)_{k\geq 1}$ and denote by $\mathbf{E}$ the expectation with respect to $\mathbf{P}$. For any $A \in C^\infty(\mathbb{R})$,*

$$\frac{d}{dt}\mathbf{E}[A\circ\pi_t] = \mathbf{E}[A'\circ\pi_t]\hat{\mu}_t(f). \tag{40}$$

Proof. Let $A \in C^\infty(\mathbb{R})$. Using the uniform Lipschitz continuity and boundedness of $(\zeta_n)_{n\geq 1}$,

$$A(\zeta_n(t+h)) - A(\zeta_n(t)) = A'(\zeta_n(t))[\zeta_n(t+h) - \zeta_n(t)] + O(h^2),$$

where the error term is uniform. Next, we integrate the above expansion with respect to μ and take $n \to \infty$ along the subsequence $(n_k)_{k\geq 1}$. Lemma 6.3 guarantees that

$$\mu[A'(\zeta_n(t))[\zeta_n(t+h) - \zeta_n(t)]] - \mu[A'(\zeta_n(t))]\mu[\zeta_n(t+h) - \zeta_n(t)] = o(1)$$

as $n \to \infty$. For the weak limit $\mathbf{P}$,

$$\lim_{k\to\infty} \mu[A(\zeta_{n_k}(t+h)) - A(\zeta_{n_k}(t))] = \mathbf{E}[A\circ\pi_{t+h} - A\circ\pi_t]$$

and

$$\lim_{k\to\infty} \mu[A'(\zeta_{n_k}(t))] = \lim_{k\to\infty} \mathbf{E}_{n_k}^\mu[A'\circ\pi_t] = \mathbf{E}[A'\circ\pi_t].$$

Recalling (29),

$$\lim_{n\to\infty} \mu[\zeta_n(t+h) - \zeta_n(t)] = \int_t^{t+h} \hat{\mu}_s(f)\, ds.$$

By Lemma 5.8,

$$\int_t^{t+h} \hat{\mu}_s(f)\, ds = \hat{\mu}_t(f)h + o(h),$$

which finishes the proof. □

https://doi.org/10.1017/etds.2016.9 Published online by Cambridge University Press

7.1. *A preliminary proposition to Theorem 3.1.* Rather than proving Theorem 3.1 directly, it is simpler to first prove the following proposition which assumes a regular initial density, recalling that $\mathbf{P}_n^\mu$ denotes the distribution of ζ_n in $C^0([0,1],\mathbb{R})$.

PROPOSITION 7.3. *Suppose f is Lipschitz continuous and μ is absolutely continuous with a strictly positive Lipschitz continuous density. The measures $\mathbf{P}_n^\mu$ converge weakly, as $n\to\infty$, to the point mass at $\zeta\in C^0([0,1],\mathbb{R})$, where*

$$\zeta(t)=\int_0^t \hat{\mu}_s(f)\,ds.$$

Proof. We show that the weak limit $\mathbf{P}$ of the subsequence $(\mathbf{P}_{n_k}^\mu)_{k\geq 1}$ is the point mass at $\zeta\in C^0([0,1],\mathbb{R})$ defined in (4). With the aid of (40), we begin by computing

$$\begin{aligned}\frac{d}{dt}\mathbf{E}[|\pi_t-\zeta(t)|^2]&=\frac{d}{dt}\mathbf{E}[\pi_t^2]-2\zeta(t)\frac{d}{dt}\mathbf{E}[\pi_t]-2\mathbf{E}[\pi_t]\frac{d}{dt}\zeta(t)+\frac{d}{dt}|\zeta(t)|^2\\&=2\mathbf{E}[\pi_t]\hat{\mu}_t(f)-2\zeta(t)\hat{\mu}_t(f)-2\mathbf{E}[\pi_t]\hat{\mu}_t(f)+2\zeta(t)\hat{\mu}_t(f)\\&=0.\end{aligned}$$

Since $\pi_0=0$ almost surely with respect to $\mathbf{P}$ and $\zeta(0)=0$, $\mathbf{E}[|\pi_t-\zeta(t)|^2]=0$ for all $t\in[0,1]$. By Tonelli's theorem,

$$\mathbf{E}\left[\int_0^1 |\pi_t-\zeta(t)|^2\,dt\right]=\int_0^1 \mathbf{E}[|\pi_t-\zeta(t)|^2]\,dt=0.$$

By the continuity of the paths, this proves the claim.

In particular, the limit $\mathbf{P}$ is independent of the initial density $\rho\in\bigcup_{L>0}\mathcal{D}_L$ and of the weakly converging subsequence $(\mathbf{P}_{n_k}^\mu)_{k\geq 1}$. Thus, we have shown that, for any initial measure μ with such a density, the sequence $(\mathbf{P}_n^\mu)_{n\geq 1}$ itself converges weakly to $\mathbf{P}$. □

7.2. *Proof of Theorem 3.1.* Next, suppose that ν is an arbitrary absolutely continuous initial probability measure with density ψ. Let $F:C^0([0,1],\mathbb{R})\to\mathbb{R}$ be an arbitrary bounded continuous function and denote $M=\sup_{\omega\in C^0([0,1],\mathbb{R})}|F(\omega)|$. By Lemma 4.5, given any $\varepsilon>0$, there exists a measure μ with density $\rho\in\bigcup_{L>0}\mathcal{D}_L$ such that

$$\|\psi-\rho\|_{L^1}\leq\frac{\varepsilon}{2M}.$$

By the established weak convergence of $(\mathbf{P}_n^\mu)_{n\geq 1}$, there exists an integer $N>0$ such that

$$|\mathbf{E}_n^\mu[F]-\mathbf{E}[F]|\leq\frac{\varepsilon}{2},\quad n\geq N.$$

Then

$$\begin{aligned}|\mathbf{E}_n^\nu[F]-\mathbf{E}_n^\mu[F]|&=\left|\int F(\zeta_n(x,\cdot))\,d\nu(x)-\int F(\zeta_n(x,\cdot))\,d\mu(x)\right|\\&\leq\int |F(\zeta_n(x,\cdot))\,(\psi(x)-\rho(x))|\,d\mathfrak{m}(x)\\&\leq M\|\psi-\rho\|_{L^1}\\&\leq\frac{\varepsilon}{2},\end{aligned}$$

https://doi.org/10.1017/etds.2016.9 Published online by Cambridge University Press

so that

$$|\mathbf{E}_n^\nu[F] - \mathbf{E}[F]| \le \varepsilon, \quad n \ge N.$$

By the portmanteau theorem, this suffices to show that $(\mathbf{P}_n^\nu)_{n\ge1}$ converges weakly to $\mathbf{P}$.

The proof of Theorem 3.1 is now complete. □

Remark 7.4. The last part of the proof implies that the results of this section hold for arbitrary absolutely continuous initial measures.

8. *Proofs of Lemma 3.5 and Theorem 3.6*

8.1. *Proof of Lemma 3.5.* If $\eta \in (0, 1)$ is arbitrary and ν is a measure having a Lipschitz continuous density ψ, then Corollary 5.7 yields

$$\begin{aligned}
&|n^{1/2}\nu(\zeta_n(\cdot\,,t)) - n^{1/2}\mathfrak{m}(\zeta_n(\cdot\,,t))| \\
&\quad\le n^{1/2}\int_0^t |\nu(f_{n,\lfloor ns\rfloor}) - \mathfrak{m}(f_{n,\lfloor ns\rfloor})|\,ds \\
&\quad\le n^{1/2}\int_0^t |\nu_{n,\lfloor ns\rfloor}(f) - \mathfrak{m}_{n,\lfloor ns\rfloor}(f)|\,ds \\
&\quad\le n^{1/2}\int_0^t \left|\int \mathcal{L}_{n,\lfloor ns\rfloor}\cdots\mathcal{L}_{n,1}(\psi-1)f\,d\mathfrak{m}\right| ds \\
&\quad\le Cn^{1/2}\int_0^t \vartheta^{ns}\,ds \le Cn^{-1/2}
\end{aligned}$$

uniformly in t. Hence the admissibility condition (6) is satisfied. This proves item (i).

Now assume that $\eta > \frac{1}{2}$ and fix $\eta' \in (\frac{1}{2}, \eta)$. Recalling (29),

$$\begin{aligned}
|n^{1/2}\zeta(t) - n^{1/2}\mathfrak{m}(\zeta_n(\cdot\,,t))| &\le n^{1/2}\int_0^t |\hat\mu_s(f) - \mathfrak{m}(f_{n,\lfloor ns\rfloor})|\,ds \\
&\le C(n^{1/2-\eta'} + n^{-1/2}\log n). \qquad (41)
\end{aligned}$$

Again, the admissibility condition (6) is satisfied. This proves item (ii).

Finally, let $\varepsilon > 0$ be arbitrary, and denote the density of ν by ψ. By Lemma 4.5, there exists a measure μ with density $\varphi \in \mathcal{D}_L$ such that $\|\varphi - \psi\|_{L^1} \le \|f\|_\infty^{-1}(\varepsilon/2)$. Then

$$\begin{aligned}
\sup_{t\in[0,1]} |\nu(\zeta_n(\cdot\,,t)) - \zeta(t)| &\le \int_0^1 |\nu(f_{n,\lfloor ns\rfloor}) - \hat\mu_s(f)|\,ds \\
&\le \frac{\varepsilon}{2} + \int_0^1 |\mu(f_{n,\lfloor ns\rfloor}) - \hat\mu_s(f)|\,ds.
\end{aligned}$$

By (29), the last term is bounded by $\varepsilon/2$ for all large enough n. This proves item (iii).

This finishes the proof of Lemma 3.5. □

8.2. *A preliminary proposition to Theorem 3.6.* Analogously to §7.1 and Theorem 3.1, in this section, we prove a simplified version of Theorem 3.6. Throughout the section we will assume that the initial measure μ is given and use the centering $c_n(t) = \mu(\zeta_n(t))$ for the process χ_n. Recall the convention from the beginning of §6 that, in this case, we write ξ_n instead of χ_n, in order to stress the explicit role of the initial measure in the centering.

https://doi.org/10.1017/etds.2016.9 Published online by Cambridge University Press

Thus ξ_n has the definition in (25) and the integral expression in (27). Each ξ_n is a random element of $C^0([0,1],\mathbb{R})$ with distribution $\mathbb{P}_n^\mu$. The expectation corresponding to $\mathbb{P}_n^\mu$ will be denoted by $\mathbb{E}_n^\mu$.

PROPOSITION 8.1. *Suppose that f is Lipschitz continuous, μ is absolutely continuous with a strictly positive Lipschitz continuous density and $c_n(t)=\mu(\zeta_n(t))$. The measures $\mathbb{P}_n^\mu$ (distribution of ξ_n) converge weakly, as $n\to\infty$, to the law of the process*

$$\chi(t)=\int_0^t \hat{\sigma}_s(f)\,dW_s.$$

Here W is a standard Brownian motion, and the stochastic integral is to be understood in the sense of Itō.

The rest of the subsection constitutes the proof of Proposition 8.1. Again, the first step is to prove tightness.

LEMMA 8.2. *Let μ be a measure with a density that is Lipschitz continuous on J_z for some $z\in\mathbb{S}$. The sequence of measures $(\mathbb{P}_n^\mu)_{n\ge1}$ is tight.*

Proof. Note that $\pi_0=0$ almost surely with respect to $\mathbb{P}_n^\mu$ for all $n\ge1$. By Kolmogorov's criterion, it is sufficient to find a constant $K>0$ such that

$$\mu[|\xi_n(t_2)-\xi_n(t_1)|^4]=\mathbb{E}_n^\mu[|\pi_{t_2}-\pi_{t_1}|^4]\le K|t_2-t_1|^2 \tag{42}$$

holds for all $t_1,t_2\in[0,1]$ and all $n\ge1$. By symmetry, the left-hand side can be expressed as

$$I=4!\,n^2\int_{t_1}^{t_2}\int_{t_1}^{s}\int_{t_1}^{r}\int_{t_1}^{u}\mu(\bar{f}_{n,\lfloor ns\rfloor}\bar{f}_{n,\lfloor nr\rfloor}\bar{f}_{n,\lfloor nu\rfloor}\bar{f}_{n,\lfloor nv\rfloor})\,dv\,du\,dr\,ds.$$

Observe that $\mu(\bar{f}_{n,\lfloor nt\rfloor})=0$ for all t. Hence, for $v<u<r<s$,

$$|\mu(\bar{f}_{n,\lfloor ns\rfloor}\bar{f}_{n,\lfloor nr\rfloor}\bar{f}_{n,\lfloor nu\rfloor}\bar{f}_{n,\lfloor nv\rfloor})|\le C\min(\vartheta^{n(s-r)},\vartheta^{n(u-v)})\le C\vartheta^{(n/2)(s-r)}\vartheta^{(n/2)(u-v)}.$$

Here the first inequality holds because Lemma 5.10 applies to both the time gap between ns and nr as well as the time gap between nu and nv, and the second inequality holds because $\min(a,b)\le a^{1/2}b^{1/2}$ for $a,b\ge0$. Assuming $t_1\le t_2$, we define $I_*=n\int_{t_1}^{t_2}\int_{t_1}^{y}\vartheta^{(n/2)(y-x)}\,dx\,dy$. Then

$$I\le 4!\,I_*^2$$

by the preceding bound. Since

$$\int_{t_1}^{y}\vartheta^{-(n/2)x}\,dx=\frac{1}{n\log\vartheta^{-1/2}}(\vartheta^{-(n/2)y}-\vartheta^{-(n/2)t_1})\le\frac{\vartheta^{-(n/2)y}}{n\log\vartheta^{-1/2}},$$

we get

$$I_*\le\frac{(t_2-t_1)}{\log\vartheta^{-1/2}},$$

which gives the desired estimate in (42). Hence, $(\mathbb{P}_n^\mu)_{n\ge1}$ is tight. □

https://doi.org/10.1017/etds.2016.9 Published online by Cambridge University Press

As in the previous section, the next step is to study the weak limit points of the sequence $(\mathbb{P}_n^\mu)_{n\geq 1}$, whose existence is guaranteed by tightness. Again, a Dynkin formula will be instrumental, albeit additional work will be required in the present setting. We write $A \in C_c^\infty(\mathbb{R})$ if $A \in C^\infty(\mathbb{R})$ and A vanishes outside a compact set.

LEMMA 8.3. *Let $\rho \in \bigcup_{L>0} \mathcal{D}_L$. Suppose that $\mathbb{P}$ is the weak limit of a subsequence $(\mathbb{P}_{n_k}^\mu)_{k\geq 1}$ and denote by $\mathbb{E}$ the expectation with respect to $\mathbb{P}$. For any $A \in C_c^\infty(\mathbb{R})$,*

$$\mathbb{E}[A \circ \pi_t] = \mathbb{E}[A \circ \pi_0] + \frac{1}{2}\int_0^t \mathbb{E}[A'' \circ \pi_s]\, \hat{\sigma}_s^2(f)\, ds. \tag{43}$$

Proof. By Taylor's theorem, there exists such a $u \in \mathbb{R}$ that

$$\begin{aligned} A(\xi_n(t+h)) = A(\xi_n(t)) + A'(\xi_n(t))[\xi_n(t+h) - \xi_n(t)] \\ + \tfrac{1}{2}A''(\xi_n(t))[\xi_n(t+h) - \xi_n(t)]^2 + \tfrac{1}{6}A'''(u)(\xi_n(t+h) - \xi_n(t))^3. \end{aligned}$$

Note that

$$\mu[|\xi_n(t+h) - \xi_n(t)|^3] \leq Ch^{2\cdot 3/4} = o(h),$$

by Jensen's inequality, together with (42). Since A''' is bounded,

$$\begin{aligned} \mu[A(\xi_n(t+h))] = \mu[A(\xi_n(t))] + \mu[A'(\xi_n(t))[\xi_n(t+h) - \xi_n(t)]] \\ + \tfrac{1}{2}\mu[A''(\xi_n(t))[\xi_n(t+h) - \xi_n(t)]^2] + o(h), \end{aligned}$$

where the error term is uniform in n and t.

Lemma 6.3 guarantees that, for $q \in \{1, 2\}$,

$$\mu[A^{(q)}(\xi_n(t))[\xi_n(t+h) - \xi_n(t)]^q] - \mu[A^{(q)}(\xi_n(t))]\mu[[\xi_n(t+h) - \xi_n(t)]^q] \to 0$$

as $n \to \infty$. By (28),

$$\mu[\xi_n(t+h) - \xi_n(t)] = 0.$$

Next, note that

$$\lim_{k\to\infty} \mu[A''(\xi_{n_k}(t))] = \mathbb{E}[A'' \circ \pi_t].$$

By Lemma 6.2,

$$\lim_{n\to\infty} \mu[[\xi_n(t+h) - \xi_n(t)]^2] = h\, \hat{\sigma}_t^2(f) + o(h).$$

Recall that the function $t \mapsto \hat{\sigma}_t^2(f)$ is continuous. Since also $t \mapsto \mathbb{E}[A'' \circ \pi_t]$ is continuous, we thus arrive at (43). □

Define the differential operator

$$\mathscr{L}_t = \frac{1}{2}\hat{\sigma}_t^2 \frac{d^2}{dx^2}.$$

Note that $\mathscr{L}_t$ appears on the right-hand side of (43). Lemma 8.3 thus leads us to conjecture that the limit process χ is a diffusion with $\mathscr{L}_t$ as its generator: that is, χ should solve the stochastic differential equation

$$d\chi(t) = \hat{\sigma}_t(f)\, dW_t, \tag{44}$$

https://doi.org/10.1017/etds.2016.9 Published online by Cambridge University Press

where W_t is a standard Brownian motion. Indeed, Itō calculus for χ, defined by (44), yields a Dynkin formula which is of precisely the same form as (43), with the law of χ in place of $\mathbb{P}$. We proceed to prove rigorously that the limit process is, indeed, characterized by (44).

Let us briefly discuss the solutions to (44), which are here always required to start at zero. Since the coefficient $\hat{\sigma}_t(f)$ is bounded in t and independent of χ, given a Brownian motion, there exists a strong solution (adapted to the filtration generated by the Brownian motion) which has continuous paths and is strongly unique (i.e. its modifications are indistinguishable). Moreover, weak solutions to (44) are unique in law; from here on we denote the associated law by Q. These facts imply that the martingale problem corresponding to the generator $\mathscr{L}_t$ and the starting point zero is well posed.

LEMMA 8.4. *The measure Q is the* unique *measure with the properties that $Q(\pi_0 = 0) = 1$ and that, for all $A \in C_c^\infty(\mathbb{R})$, the process*

$$M_t = A \circ \pi_t - A \circ \pi_0 - \int_0^t \mathscr{L}_s A \circ \pi_s \, ds, \quad t \in [0, 1], \tag{45}$$

is a martingale with respect to Q and the filtration $(\mathfrak{F}_t)_{0 \le t \le 1}$, where $\mathfrak{F}_t$ is the sigma-algebra on $C^0([0, 1], \mathbb{R})$ generated by $\{\pi_s : 0 \le s \le t\}$.

We refer to [**30**] for the proofs of the above statements. Also [**13**] is a helpful text on stochastic analysis.

The next result states that also the measure $\mathbb{P}$ solves the above martingale problem. Hence, it follows, from the above lemma, that $\mathbb{P} = Q$.

PROPOSITION 8.5. *Suppose the density of the initial measure μ is in $\bigcup_{L>0} \mathcal{D}_L$ and that $\mathbb{P}$ is the weak limit of a subsequence $(\mathbb{P}_{n_k})_{k \ge 1}$. Then, given any $A \in C_c^\infty(\mathbb{R})$, the process $(M_t)_{t \in [0,1]}$ defined in (45) is a martingale with respect to $\mathbb{P}$ and the filtration $(\mathfrak{F}_t)_{0 \le t \le 1}$. In particular,*

$$\mathbb{P} = Q.$$

Note that the limit $\mathbb{P}$ is then independent of the initial density $\rho \in \bigcup_{L>0} \mathcal{D}_L$ and of the weakly converging subsequence $(\mathbb{P}_{n_k}^\mu)_{k \ge 1}$. Thus, for any initial measure μ with such a density, the sequence $(\mathbb{P}_n^\mu)_{n \ge 1}$ itself converges weakly to $\mathbb{P}$. Accordingly, we have identified the limit of ξ_n to be the stochastic process χ appearing in (44). In particular, once this proposition has been proved, we will have proved Proposition 8.1.

Proof of Proposition 8.5. Let $L > 0$ and $z \in \mathbb{S}$ be such that the logarithm of the initial density ρ is Lipschitz continuous on J_z with constant L. Since A and A'' are bounded, $\mathbb{E}[|M_t|] < \infty$ for all $t \in [0, 1]$. It remains to prove that

$$\mathbb{E}[M_t - M_r \mid \mathfrak{F}_r] = 0$$

whenever $0 \le r \le t \le 1$. The martingale condition above is equivalent to

$$\mathbb{E}[B_1 \circ \pi_{t_1} \cdots B_m \circ \pi_{t_m} (M_t - M_r)] = 0 \tag{46}$$

whenever $m \ge 1$, $B_1, \ldots, B_m : \mathbb{R} \to \mathbb{R}$ are bounded, Lipschitz continuous functions and $0 < t_1 < \cdots < t_m \le r < t \le 1$. We now fix such numbers and functions for good.

https://doi.org/10.1017/etds.2016.9 Published online by Cambridge University Press

Let us fix $q \in (0, \frac{1}{2})$ and write $K_n = \lfloor n^q(t-r) \rfloor$ and $\delta_n = (t-r)/K_n$. Since

$$M_t - M_r = \sum_{k=0}^{K_n-1} (M_{r+(k+1)\delta_n} - M_{r+k\delta_n}),$$

equation (46) will follow once we establish that

$$K_n \sup_{r \le u \le t-\delta_n} \mu[B_1(\xi_n(t_1)) \cdots B_m(\xi_n(t_m))(M_{u+\delta_n} - M_u)(\xi_n)] = o(1),$$

as $n \to \infty$.

We proceed as in the proof of Lemma 6.3, resorting to the induced partition $\mathcal{P}_{z,n,u} = \{I_{z,n,u,j}\}_{j=1}^{N_{n,u}}$ of the arc $J_z = \mathbb{S}\backslash\{z\}$. Let $\hat{x}_{z,n,u,j}$ denote the midpoint of $I_{z,n,u,j}$ and $c_{z,n,u,j} = \xi_n(\hat{x}_{z,n,u,j}, u)$ the value of $\xi_n(\,\cdot\,, u)$ at the midpoint. It will then be convenient to define the function $\hat{\xi}_{n,u} : \mathbb{S} \times [0,1] \to \mathbb{R}$ by setting

$$\hat{\xi}_{n,u}(x,t) = \xi_n(x,t) - \xi_n(x,u) + c_{z,n,u,j}$$

for all $x \in I_{z,n,u,j}$ and all j. (For brevity, we suppress the z-dependence of $\hat{\xi}_{n,u}$.) We think of $\hat{\xi}_{n,u}$ as a modification of ξ_n according to the value of the latter process at time u. By (38),

$$\sup_{x\in\mathbb{S}} |\xi_n(x,t) - \hat{\xi}_{n,u}(x,t)| = \max_j \sup_{x \in I_{z,n,u,j}} |\xi_n(x,u) - \xi_n(\hat{x}_{z,n,u,j}, u)| \le Cn^{-1/2}$$

uniformly in u. Since A and A'' are Lipschitz continuous, the functional $M_{u+\delta_n} - M_u : C^0([0,1], \mathbb{R}) \to \mathbb{R}$ satisfies

$$\sup_{x\in\mathbb{S}} |(M_{u+\delta_n} - M_u)(\xi_n(x,\cdot)) - (M_{u+\delta_n} - M_u)(\hat{\xi}_{n,u}(x,\cdot))| \le Cn^{-1/2}.$$

Because $B_1, \ldots, B_m$ are, furthermore, bounded, Lipschitz continous functions, (39) yields

$$\begin{aligned}
&\mu[B_1(\xi_n(t_1)) \cdots B_m(\xi_n(t_m))(M_{u+\delta_n} - M_u)(\xi_n)] \\
&\quad = \mu[B_1(\xi_n(t_1)) \cdots B_m(\xi_n(t_m))(M_{u+\delta_n} - M_u)(\hat{\xi}_{n,u})] + O(n^{-1/2}) \\
&\quad = \sum_j \mu[1_{I_{z,n,u,j}} B_1(\xi_n(t_1)) \cdots B_m(\xi_n(t_m))(M_{u+\delta_n} - M_u)(\hat{\xi}_{n,u})] + O(n^{-1/2}) \\
&\quad = \sum_j \mu_{z,n,u,j}[B_1(\xi_n(t_1)) \cdots B_m(\xi_n(t_m))]\mu[1_{I_{z,n,u,j}}(M_{u+\delta_n} - M_u)(\hat{\xi}_{n,u})] \\
&\qquad + O(n^{-1/2}) \\
&\quad = \sum_j \mu[1_{I_{z,n,u,j}} B_1(\xi_n(t_1)) \cdots B_m(\xi_n(t_m))]\mu_{z,n,u,j}[(M_{u+\delta_n} - M_u)(\hat{\xi}_{n,u})] \\
&\qquad + O(n^{-1/2}).
\end{aligned}$$

Here the error term is, again, uniform in u. Since $K_n = o(n^{1/2})$, it thus suffices to show that

$$K_n \sup_{r \le u \le t-\delta_n} \max_j |\mu_{z,n,u,j}[(M_{u+\delta_n} - M_u)(\hat{\xi}_{n,u})]| = o(1).$$

Here

$$(M_{u+\delta_n} - M_u)(\hat{\xi}_{n,u}) = A(\hat{\xi}_{n,u}(u+\delta_n)) - A(\hat{\xi}_{n,u}(u)) - \int_u^{u+\delta_n} \mathscr{L}_s A(\hat{\xi}_{n,u}(s))\, ds.$$

https://doi.org/10.1017/etds.2016.9 Published online by Cambridge University Press

For $x \in I_{z,n,u,j}$, we Taylor expand $A(\hat{\xi}_{n,u}(x, u+\delta_n))$ at $\hat{\xi}_{n,u}(x,u) = c_{z,n,u,j}$. By Taylor's theorem, there exists $\kappa_{z,n,u,j}(x) \in \mathbb{R}$ such that

$$\begin{aligned}
&(M_{u+\delta_n} - M_u)(\hat{\xi}_{n,u}) \\
&\quad = A'(c_{z,n,u,j})[\xi_n(u+\delta_n) - \xi_n(u)] \\
&\qquad + \left[\frac{1}{2}A''(c_{z,n,u,j})\,[\xi_n(u+\delta_n) - \xi_n(u)]^2 - \int_u^{u+\delta_n} \mathscr{L}_s A(\hat{\xi}_{n,u}(s))\,ds\right] \\
&\qquad + \frac{1}{6}A'''(\kappa_{z,n,u,j})[\xi_n(u+\delta_n) - \xi_n(u)]^3.
\end{aligned}$$

We bound the term on each line on the right-hand side separately. The second line requires demonstrating cancelations in the difference and is saved for last.

It is instrumental that (assuming $\lfloor nt_1 \rfloor \geq N(L)$) the density $\rho_{z,n,u,j}$ of $\mu_{z,n,u,j}$ satisfies

$$\tilde{\rho}_{z,n,u,j} = \mathcal{L}_{n,\lfloor nu \rfloor} \cdots \mathcal{L}_{n,1}\rho_{z,n,u,j} \in \mathcal{D}_{L_*}$$

by part (iii) of Lemma 5.2. We denote the measure corresponding to $\tilde{\rho}_{z,n,u,j}$ by $\tilde{\mu}_{z,n,u,j}$. Moreover, it will be convenient to define

$$\xi_n^{(u)}(h) = n^{1/2}\int_u^{u+h} f_{n,\lfloor ns \rfloor,\lfloor nu \rfloor+1} - \mu_{n,\lfloor nu \rfloor}(f_{n,\lfloor ns \rfloor,\lfloor nu \rfloor+1})\,ds,$$

where $f_{n,k,l} = f \circ T_{n,k} \circ \cdots \circ T_{n,l}$ for $k \geq l$ and $f_{n,k,k+1} = f$. Then

$$\xi_n(u+h) - \xi_n(u) = \xi_n^{(u)}(h) \circ T_{n,\lfloor nu \rfloor} \circ \cdots \circ T_{n,1}.$$

It is helpful to think of $\xi_n^{(u)}$ as ξ_n shifted along the curve γ. Below, we will need to change the centering of $\xi_n^{(u)}$, so we now define

$$\xi_n^{\nu,(u)}(h) = n^{1/2}\int_u^{u+h} f_{n,\lfloor ns \rfloor,\lfloor nu \rfloor+1} - \nu(f_{n,\lfloor ns \rfloor,\lfloor nu \rfloor+1})\,ds = \xi_n^{(u)}(h) - \nu\big[\xi_n^{(u)}(h)\big]$$

for an arbitrary measure ν with density $\psi \in \mathcal{D}_{L_*}$. By Corollary 5.7 applied to the difference $g = \psi - \rho_{n,\lfloor nu \rfloor}$, we have the uniform bound

$$|\nu(f_{n,\lfloor ns \rfloor,\lfloor nu \rfloor+1}) - \mu_{n,\lfloor nu \rfloor}(f_{n,\lfloor ns \rfloor,\lfloor nu \rfloor+1})| = O(\vartheta^{\lfloor ns \rfloor - \lfloor nu \rfloor}) = O(\vartheta^{n(s-u)}).$$

Hence, an integration yields

$$\nu\big[\xi_n^{(u)}(h)\big] = n^{1/2}\int_u^{u+h} \nu(f_{n,\lfloor ns \rfloor,\lfloor nu \rfloor+1}) - \mu_{n,\lfloor nu \rfloor}(f_{n,\lfloor ns \rfloor,\lfloor nu \rfloor+1})\,ds = O(n^{-1/2})$$

uniformly in u, h and ν, so that

$$\xi_n^{\nu,(u)}(h) = \xi_n^{(u)}(h) + O(n^{-1/2}) \tag{47}$$

and, by $\nu\big[\xi_n^{\nu,(u)}(h)\big] = 0$,

$$\nu[\xi_n^{(u)}(h)] = O(n^{-1/2}) \tag{48}$$

uniformly in u, h and ν. Note that the error terms above are independent of x.

The first term. Since

$$\mu_{z,n,u,j}[A'(c_{z,n,u,j})[\xi_n(u+\delta_n) - \xi_n(u)]] = A'(c_{z,n,u,j})\mu_{z,n,u,j}[\xi_n(u+\delta_n) - \xi_n(u)],$$

https://doi.org/10.1017/etds.2016.9 Published online by Cambridge University Press

where A' is bounded, we only need a bound on the second factor on the right. Here

$$\mu_{z,n,u,j}[\xi_n(u+\delta_n)-\xi_n(u)] = \tilde{\mu}_{z,n,u,j}[\xi_n^{(u)}(\delta_n)].$$

Recalling the earlier remark on $\tilde{\mu}_{z,n,u,j}$, (48) yields

$$\tilde{\mu}_{z,n,u,j}[\xi_n^{(u)}(\delta_n)] = o(K_n^{-1}).$$

In particular,

$$K_n \sup_{r\le u\le t-\delta_n} \max_j |\mu_{z,n,u,j}[A'(c_{z,n,u,j})[\xi_n(u+\delta_n)-\xi_n(u)]]| = o(1).$$

The third term. Using the boundedness of A''', together with Jensen's inequality,

$$|\mu_{z,n,u,j}[A'''(\kappa_{z,n,u,j})[\xi_n(u+\delta_n)-\xi_n(u)]^3]| \le C\mu_{z,n,u,j}[[\xi_n(u+\delta_n)-\xi_n(u)]^4]^{3/4}.$$

Here

$$\mu_{z,n,u,j}[[\xi_n(u+\delta_n)-\xi_n(u)]^4] = \tilde{\mu}_{z,n,u,j}[\xi_n^{(u)}(\delta_n)^4].$$

Let us consider an arbitrary measure ν with density $\psi \in \mathcal{D}_{L_*}$, instead of $\tilde{\mu}_{z,n,u,j}$. First, we have the uniform bound

$$\nu[\xi_n^{\nu,(u)}(h)^4] = O(h^2),$$

which is analogous to (42). Together with (47), Jensen's inequality then shows that

$$\nu[\xi_n^{(u)}(h)^4] = O(h^2 + h^{3/2}n^{-1/2} + hn^{-1} + n^{-2}) \tag{49}$$

uniformly. In particular,

$$\tilde{\mu}_{z,n,u,j}[|\xi_n^{(u)}(\delta_n)|^3] = O(\delta_n^2 + \delta_n^{3/2}n^{-1/2} + \delta_n n^{-1} + n^{-2})^{3/4} = o(K_n^{-1}),$$

so that

$$K_n \sup_{r\le u\le t-\delta_n} \max_j |\mu_{z,n,u,j}[A'''(\kappa_{z,n,u,j})[\xi_n(u+\delta_n)-\xi_n(u)]^3]| = o(1).$$

The second term. Note that

$$\begin{aligned}&\mu_{z,n,u,j}[\tfrac{1}{2}A''(c_{z,n,u,j})[\xi_n(u+\delta_n)-\xi_n(u)]^2]\\ &\quad = \tfrac{1}{2}A''(c_{z,n,u,j})\mu_{z,n,u,j}[[\xi_n(u+\delta_n)-\xi_n(u)]^2],\end{aligned}$$

where

$$\mu_{z,n,u,j}[[\xi_n(u+\delta_n)-\xi_n(u)]^2] = \tilde{\mu}_{z,n,u,j}[\xi_n^{(u)}(\delta_n)^2].$$

On the other hand,

$$\begin{aligned}&\mu_{z,n,u,j}\left[\int_u^{u+\delta_n} \mathscr{L}_s A(\hat{\xi}_{n,u}(s))\,ds\right]\\ &\quad = \int_u^{u+\delta_n} \mu_{z,n,u,j}\left[\frac{1}{2}\hat{\sigma}_s^2(f)A''(\hat{\xi}_{n,u}(s))\right]ds\\ &\quad = \frac{1}{2}\int_u^{u+\delta_n} \hat{\sigma}_s^2(f)\mu_{z,n,u,j}[A''(\xi_n(s)-\xi_n(u)+c_{z,n,u,j})]\,ds\\ &\quad = \frac{1}{2}\int_u^{u+\delta_n} \hat{\sigma}_s^2(f)\mu_{z,n,u,j}[A''(\xi_n^{(u)}(s-u)\circ T_{n,\lfloor nu\rfloor}\circ\cdots\circ T_{n,1}+c_{z,n,u,j})]\,ds\\ &\quad = \frac{1}{2}\int_u^{u+\delta_n} \hat{\sigma}_s^2(f)\tilde{\mu}_{z,n,u,j}[A''(\xi_n^{(u)}(s-u)+c_{z,n,u,j})]\,ds.\end{aligned}$$

https://doi.org/10.1017/etds.2016.9 Published online by Cambridge University Press

By Taylor's theorem, there exists $\tilde{\kappa}_{z,n,u,j,s}(x) \in \mathbb{R}$, such that

$$\tilde{\mu}_{z,n,u,j}[A''(\xi_n^{(u)}(s-u) + c_{z,n,u,j})] = A''(c_{z,n,u,j}) + \tilde{\mu}_{z,n,u,j}[A'''(\tilde{\kappa}_{z,n,u,j,s})\xi_n^{(u)}(s-u)].$$

Using the boundedness of A''', together with Jensen's inequality, (49) yields

$$|\tilde{\mu}_{z,n,u,j}[A'''(\tilde{\kappa}_{z,n,u,j,s})\xi_n^{(u)}(s-u)]| \le C\tilde{\mu}_{z,n,u,j}[\xi_n^{(u)}(s-u)^4]^{1/4} = O(\delta_n^{1/2}).$$

Accordingly,

$$\begin{aligned}\mu_{z,n,u,j}\left[\int_u^{u+\delta_n} \mathscr{L}_s A(\hat{\xi}_{n,u}(s))\, ds\right] &= \frac{1}{2}A''(c_{z,n,u,j})\int_u^{u+\delta_n} \hat{\sigma}_s^2(f)\, ds + O(\delta_n^{3/2}) \\ &= \frac{1}{2}A''(c_{z,n,u,j})\hat{\sigma}_u^2(f)\delta_n + o(\delta_n).\end{aligned}$$

In the second line, we used (32). We remark that the error term $o(\delta_n) = o(K_n^{-1})$ is uniform in u and j and that A'' is bounded. In order to prove that

$$K_n \sup_{r\le u\le t-\delta_n} \max_j \left|\mu_{z,n,u,j}\left[\frac{1}{2}A''(c_{z,n,u,j})[\xi_n(u+\delta_n) - \xi_n(u)]^2 - \int_u^{u+\delta_n} \mathscr{L}_s A(\hat{\xi}_{n,u}(s))\, ds\right]\right|$$

tends to zero as $n \to \infty$, it thus only remains to show that

$$K_n \sup_{r\le u\le t-\delta_n} \max_j |\tilde{\mu}_{z,n,u,j}[\xi_n^{(u)}(\delta_n)^2] - \hat{\sigma}_u^2(f)\delta_n| = o(1).$$

Let us, again, consider an arbitrary measure ν with density $\psi \in \mathcal{D}_{L_*}$, instead of $\tilde{\mu}_{z,n,u,j}$. Recalling (47) and $\nu\big[\xi_n^{\nu,(u)}(h)\big] = 0$, we get the uniform estimate

$$\nu[\xi_n^{(u)}(\delta_n)^2] = \nu[\xi_n^{\nu,(u)}(\delta_n)^2] + O(n^{-1}).$$

Analogously to Lemma 6.2,

$$K_n|\nu[\xi_n^{\nu,(u)}(\delta_n)^2] - \hat{\sigma}_u^2(f)\delta_n| = K_n|o(\delta_n) + o(n^{-1/2})| = o(1)$$

uniformly in u and ν. The last bounds, combined, yield the desired bound.

This finishes the proof of Proposition 8.5. □

The proof of Proposition 8.1 is now complete. □

8.3. *Finishing the proof of Theorem 3.6.* It remains to upgrade Proposition 8.1 to the full version of Theorem 3.6. The upgrade entails relaxing the regularity assumption on the initial measure as well as the assumption that the centering sequence be defined in terms of the initial measure. To facilitate these changes, let us introduce the explicit notation

$$\chi_n^{\nu}(t) = n^{1/2}\zeta_n(t) - n^{1/2}\nu(\zeta_n(t)),$$

for any measure ν, and

$$\chi_n^{c_n}(t) = n^{1/2}\zeta_n(t) - n^{1/2}c_n(t),$$

for any centering sequence $(c_n)_{n\ge1}$. Note that all functions above, including c_n, depend on x, but, following our earlier convention, we suppress it from the notation. Given an initial measure μ, we denote the laws of χ_n^{ν} and $\chi_n^{c_n}$ by $\mathbb{P}_n^{\mu,\nu}$ and $\mathbb{P}_n^{\mu,c_n}$, respectively. The respective expectations are denoted $\mathbb{E}_n^{\mu,\nu}$ and $\mathbb{E}_n^{\mu,c_n}$.

https://doi.org/10.1017/etds.2016.9 Published online by Cambridge University Press

Assume that μ is an arbitrary absolutely continuous measure with density ρ and $(c_n)_{n\geq 1}$ is an arbitrary centering sequence admissible with respect to μ (see Definition 3.3). Our proof of Theorem 3.6 amounts to showing that there exists a measure ν with density $\psi \in \bigcup_{L>0} \mathcal{D}_L$ with the following properties: (i) $\mathbb{P}_n^{\nu,\nu}$ approximates $\mathbb{P}$, (ii) $\mathbb{P}_n^{\nu,\mathfrak{m}}$ approximates $\mathbb{P}_n^{\nu,\nu}$, (iii) $\mathbb{P}_n^{\mu,\mathfrak{m}}$ approximates $\mathbb{P}_n^{\nu,\mathfrak{m}}$ and (iv) $\mathbb{P}_n^{\mu,c_n}$ approximates $\mathbb{P}_n^{\mu,\mathfrak{m}}$ arbitrarily well for all large enough $n \geq 1$. This will be accomplished using the portmanteau theorem, as follows.

Let $F : C^0([0,1],\mathbb{R}) \to \mathbb{R}$ be an arbitrary bounded Lipschitz continuous function and $\varepsilon > 0$ an arbitrary number. Denote $M = \sup_{\omega\in C^0([0,1],\mathbb{R})} |F(\omega)|$ and $\ell = \mathrm{Lip}(F)$.

Step (i). Observe that Proposition 8.1 applies directly to $\mathbb{P}_n^{\nu,\nu}$: for any $\psi \in \bigcup_{L>0} \mathcal{D}_L$, there exists an integer $N_1 > 0$ such that

$$|\mathbb{E}[F] - \mathbb{E}_n^{\nu,\nu}[F]| \leq \frac{\varepsilon}{4}, \quad n \geq N_1.$$

Step (ii). Since $\psi \in \bigcup_{L>0} \mathcal{D}_L$, the centering sequence $\nu(\zeta_n(t))$ is admissible with respect to ν (see Lemma 3.5). By Remark 3.4, there exists an integer $N_2 > 0$ such that $|n^{1/2}\nu(\zeta_n(t)) - n^{1/2}\mathfrak{m}(\zeta_n(t))| \leq \varepsilon/4\ell$ for all $n \geq N_2$. Thus,

$$\begin{aligned} |\mathbb{E}_n^{\nu,\nu}[F] - \mathbb{E}_n^{\nu,\mathfrak{m}}[F]| &= \left| \int F(\chi_n^{\nu}(x,\cdot)) - F(\chi_n^{\mathfrak{m}}(x,\cdot))\, d\nu(x) \right| \\ &\leq \ell \frac{\varepsilon}{4\ell} \\ &\leq \frac{\varepsilon}{4}, \quad n \geq N_2. \end{aligned}$$

Step (iii). By Lemma 4.5, we may assume that $\psi \in \bigcup_{L>0} \mathcal{D}_L$ satisfies

$$\|\psi - \rho\|_{L^1} \leq \frac{\varepsilon}{4M}.$$

Then

$$\begin{aligned} |\mathbb{E}_n^{\nu,\mathfrak{m}}[F] - \mathbb{E}_n^{\mu,\mathfrak{m}}[F]| &= \left| \int F(\chi_n^{\mathfrak{m}}(x,\cdot))\, d\nu(x) - \int F(\chi_n^{\mathfrak{m}}(x,\cdot))\, d\mu(x) \right| \\ &\leq \int |F(\chi_n^{\mathfrak{m}}(x,\cdot))\, (\psi - \rho)(x)|\, d\mathfrak{m}(x) \\ &\leq M\|\psi - \rho\|_{L^1} \\ &\leq \frac{\varepsilon}{4}, \quad n \geq 1. \end{aligned}$$

Step (iv). Recall that the centering sequence c_n is assumed admissible with respect to μ. Let us denote $E_n = \{x \in \mathbb{S} : \sup_{t\in[0,1]} |n^{1/2}c_n(x,t) - n^{1/2}\mathfrak{m}(\zeta_n(t))| > \varepsilon/8\ell\}$. There exists an integer $N_3 > 0$ such that $\mu(E_n) < \varepsilon/16M$ for all $n \geq N_3$. Splitting $\int = \int_{E_n} + \int_{\mathbb{S}\setminus E_n}$,

$$\begin{aligned} |\mathbb{E}_n^{\mu,\mathfrak{m}}[F] - \mathbb{E}_n^{\mu,c_n}[F]| &= \left| \int F(\chi_n^{\mathfrak{m}}(x,\cdot)) - F(\chi_n^{c_n}(x,\cdot))\, d\mu(x) \right| \\ &\leq 2M\mu(E_n) + \ell\frac{\varepsilon}{8\ell}(1 - \mu(E_n)) \\ &\leq \frac{\varepsilon}{4}, \quad n \geq N_3. \end{aligned}$$

https://doi.org/10.1017/etds.2016.9 Published online by Cambridge University Press

Collecting the bounds,

$$|\mathbb{E}[F] - \mathbb{E}_n^{\mu,c_n}[F]| \le \varepsilon, \quad n \ge \max(N_1, N_2, N_3).$$

By the portmanteau theorem, this suffices to show that $(\mathbb{P}_n^{\mu,c_n})_{n\ge1}$ converges weakly to $\mathbb{P}$.

The proof of Theorem 3.6 is now complete. □

Remark 8.6. The last part of the proof implies that the results of this section hold for arbitrary absolutely continuous initial measures and admissible centering sequences.

9. *Proofs of the generalizations*

As mentioned at the beginning, little in the proofs of Theorems 3.1 and 3.6 changes when one passes to their generalizations, Theorems 3.9 and 3.10. In order to keep the presentation as lucid as possible, we have elected to save the generalizations for last. Here we expect the reader to be well familiar with all the preceding sections. Indeed, we will only point the reader to the straightforward adjustments required there to complete the proofs.

9.1. *Proof of Theorem 3.9.* The first changes occur in §6. In Lemma 6.1, $\hat\sigma_t^2$ is now the $d \times d$ matrix defined in Theorem 3.9. Working with vector components, it amounts to a minor modification of the proof of Lemma 6.1 to show that

$$\begin{aligned}\hat\sigma_t^2(f) &= \hat\mu_t[\hat f_t \otimes \hat f_t] + \sum_{k=1}^{\infty} \hat\mu_t[\hat f_t \otimes (\hat f_t \circ \gamma_t^k) + (\hat f_t \circ \gamma_t^k) \otimes \hat f_t] \\ &= \hat\mu_t[\hat f_t \otimes \hat f_t] + \sum_{k=1}^{\infty} \mathfrak{m}[\hat f_t \otimes \mathcal{L}_t^k(\hat\rho_t \hat f_t) + (\mathcal{L}_t^k(\hat\rho_t \hat f_t)) \otimes \hat f_t], \end{aligned} \tag{50}$$

and that the dependence on t is continuous. Similarly, we see that Lemma 6.2 remains true, except that we have the $d \times d$ matrix $\mu[[\xi_n(t) - \xi_n(s)] \otimes [\xi_n(t) - \xi_n(s)]]$ in place of the scalar $\mu[[\xi_n(t) - \xi_n(s)]^2]$. In (39), the functions B_i are now bounded and Lipschitz continuous from $\mathbb{R}^d$ to $\mathbb{R}$, but the bound remains true. Lemma 6.3 is modified as follows: we need $A \in C^\infty(\mathbb{R}^d, \mathbb{R})$ in part (i) and $A \in C_c^\infty(\mathbb{R}^d, \mathbb{R})$ in part (ii). Then (i) and (ii) with $q = 1$ continue to hold componentwise. In the case $q = 2$, the expression $[\xi_n(t) - \xi_n(s)]^2$ is replaced by $[\xi_n(t) - \xi_n(s)] \otimes [\xi_n(t) - \xi_n(s)]$. The proof remains identical.

Let us proceed to §7. In Lemma 7.2, we get the Dynkin formula

$$\frac{d}{dt}\mathbf{E}[A \circ \pi_t] = \mathbf{E}[\nabla A \circ \pi_t] \cdot \hat\mu_t(f)$$

in place of (40), for any $A \in C^\infty(\mathbb{R}^d, \mathbb{R})$. Indeed, the modification of Lemma 6.3 above implies that

$$\mu[\nabla A(\zeta_n(t)) \cdot [\zeta_n(t+h) - \zeta_n(t)]] - \mu[\nabla A(\zeta_n(t))] \cdot \mu[\zeta_n(t+h) - \zeta_n(t)] = o(1)$$

as $n \to \infty$; otherwise the proof is similar. The proof of Proposition 7.3 remains the same up to passing to vector notation.

Finally, let us point out the changes in §8. In the tightness proof of Lemma 8.2, we have the bound in (42) separately for each vector component $\xi_n^{(i)}$, $1 \le i \le d$. The vector-valued

https://doi.org/10.1017/etds.2016.9 Published online by Cambridge University Press

case follows by an application of the Cauchy–Schwarz inequality. In Lemma 8.3, we get, for any $A \in C_c^\infty(\mathbb{R}^d, \mathbb{R})$, the Dynkin formula

$$\mathbb{E}[A \circ \pi_t] = \mathbb{E}[A \circ \pi_0] + \frac{1}{2} \sum_{i,j=1}^{d} \int_0^t \mathbb{E}[\partial_{ij}^2 A \circ \pi_s] \hat{\sigma}_s^2(f)_{ij}\, ds, \tag{51}$$

in place of (43). This identity is obtained as before, by Taylor expanding $\mu[A(\xi_n(t))]$ and using the modifications of Lemmas 6.3 and 6.2 above. Let us define $\hat{\sigma}_t(f)$ as the square root of the $d \times d$ covariance matrix $\hat{\sigma}_t^2(f)$. Then the stochastic differential equation (44), where W_t is an $\mathbb{R}^d$-valued standard Brownian motion, has the partial differential operator

$$\mathscr{L}_t = \frac{1}{2} \sum_{i,j=1}^{d} \hat{\sigma}_t^2(f)_{ij} \frac{\partial^2}{\partial x_i\, \partial x_j}$$

as its generator, so that the Dynkin formula for χ has the same form as (51). Again, the martingale problem corresponding to the generator $\mathscr{L}_t$ and the starting point zero is well posed (see Lemma 8.4 with the change $A \in C_c^\infty(\mathbb{R}^d, \mathbb{R})$). We refer to [**30**] for the proofs of these statements. Given the above changes, the proof of the martingale property in Proposition 8.5 generalizes to the vector-valued case in a straightforward manner, which shows that the limit law $\mathbb{P}$ of ξ_n is the one of the process χ appearing in (44). To complete the proof, note that §8.3 continues to apply, *mutatis mutandis*, switching from the space $C^0([0, 1], \mathbb{R})$ of real-valued functions to $C^0([0, 1], \mathbb{R}^d)$. □

9.2. *Proof of Theorem 3.10.* Let $0 = \hat{\tau}_0 < \hat{\tau}_1 < \cdots < \hat{\tau}_{m-1} < \hat{\tau}_m = 1$ be the endpoints of the partition elements I_i, $1 \le i \le m$.

The jumps in the curve γ cause certain estimates in §5 to hold only piecewise, on each partition element. The first change occurs in Lemma 5.5: (18) continues to hold if s and t are in the same partition element. Likewise, (22) in Lemma 5.8 continues to hold in the same piecewise sense. Another change occurs in Lemma 5.9, where the condition on k is replaced by the piecewise analogue $n\hat{\tau}_i + b \log n \le k < n\hat{\tau}_{i+1}$ for some i. In Lemma 5.9 the condition on t is replaced by $\hat{\tau}_i + bn^{-1} \log n \le t < \hat{\tau}_{i+1}$ and $s \in I_i$ for some i. Otherwise §5 remains intact. To recapitulate, the essential change is that the difference $\rho_{n,\lfloor nt \rfloor} - \hat{\rho}_t$ does not remain small when t passes one of the finite number of singularities $\hat{\tau}_i$; to regain smallness, it is necessary to wait another $b \log n$ iterates exactly as was the case at $t = 0$ before. Moreover, $\hat{\rho}_t - \hat{\rho}_s$ is only small if t and s are in the same partition element and close to each other. However, it is important to point out that the singularities do not affect the *regularity bounds* of the densities $\rho_{n,\lfloor nt \rfloor}$ in any way.

In §6 there are similar changes: (29) holds if $\hat{\tau}_i + bn^{-1} \log n \le s < \hat{\tau}_{i+1}$ for some i. We also note that the dependence of $\hat{\sigma}_t^2(f)$ (and of $\hat{\mu}_t(f)$) on t is piecewise-continuous, which is inconsequential (see the proof of Lemma 6.1 (and of Lemma 5.5)). Lemma 6.2 remains intact: in order to avoid singularities, one excludes, in the domain of the s-integral on the right-hand side of (35), a neighborhood of radius $2a_n$ centered at each singularity $\hat{\tau}_i$, which only results in another error of the same order na_n^2, as before. These are the only parts in §6 that require attention.

The above changes in no way affect §7, so we are left with §8. There the proof of Lemma 3.5 stands; in the integral in (41) one has to take into account the above change in

https://doi.org/10.1017/etds.2016.9 Published online by Cambridge University Press

the condition of (29) and thus remove an interval of length $bn^{-1}\log n$ at each singularity $\hat{\tau}_i$, but this does not affect the error term. The above changes do not affect the rest of §8.

The generalization to vector-valued observables is now achieved exactly as in the proof of Theorem 3.9. □

Acknowledgements. This work was supported partially by the Academy of Finland. Mikko Stenlund acknowledges the ERC grant MPOES, Jane and Aatos Erkko Foundation, and Emil Aaltosen Säätiö. Neil Dobbs was supported by the ERC Bridges project. The authors are grateful to Carlangelo Liverani for helpful correspondence and to Dario Gasbarra and Kalle Kytölä for useful remarks.

References

[**1**] R. Aimino, H. Hu, M. Nicol, A. Török and S. Vaienti. Polynomial loss of memory for maps of the interval with a neutral fixed point. *Discrete Contin. Dyn. Syst.* **35**(3) (2015), 793–806.
[**2**] V. I. Bakhtin. Random processes generated by a hyperbolic sequence of mappings. I. *Izv. Ross. Akad. Nauk Ser. Mat.* **58**(2) (1994), 40–72.
[**3**] V. Baladi. *Positive Transfer Operators and Decay of Correlations (Advanced Series in Nonlinear Dynamics, 16)*. World Scientific Publishing Co., Inc., River Edge, NJ, 2000.
[**4**] P. H. Baxendale. Stability and equilibrium properties of stochastic flows of diffeomorphisms. *Diffusion Processes and Related Problems in Analysis,* Vol. II *(Charlotte, NC, 1990) (Progress in Probability, 27)*. Birkhäuser Boston, Boston, MA, 1992, pp. 3–35.
[**5**] P. Billingsley. Convergence of probability measures. *Wiley Series in Probability and Statistics: Probability and Statistics*, 2nd edn. John Wiley & Sons, Inc., New York, 1999, A Wiley-Interscience Publication.
[**6**] X. Bressaud and C. Liverani. Anosov diffeomorphisms and coupling. *Ergod. Th. & Dynam. Sys.* **22**(1) (2002), 129–152.
[**7**] N. Chernov. Advanced statistical properties of dispersing billiards. *J. Stat. Phys.* **122**(6) (2006), 1061–1094.
[**8**] N. Chernov and D. Dolgopyat. Brownian Brownian motion. I. *Mem. Amer. Math. Soc.* **198**(927) (2009), viii+193.
[**9**] J.-P. Conze and A. Raugi. Limit theorems for sequential expanding dynamical systems on [0, 1]. *Ergodic Theory and Related Fields (Contemporary Mathematics, 430)*. American Mathematical Society, Providence, RI, 2007, pp. 89–121.
[**10**] J. De Simoi and C. Liverani. The martingale approach after Varadhan and Dolgopyat. *Hyperbolic Dynamics, Fluctuations and Large Deviations (Proceedings of Symposia in Pure Mathematics, 89)*. American Mathematical Society, Providence, RI, 2015, pp. 311–339.
[**11**] D. Dolgopyat. Averaging and invariant measures. *Mosc. Math. J.* **5**(3) (2005), 537–576, 742.
[**12**] A. P. Douce. *Thermodynamics of the Earth and Planets*. Cambridge University Press, Cambridge, 2011.
[**13**] R. Durrett. *Stochastic Calculus: A Practical Introduction (Probability and Stochastics Series)*. CRC Press, Boca Raton, FL, 1996.
[**14**] S. Gouëzel and C. Liverani. Banach spaces adapted to Anosov systems. *Ergod. Th. & Dynam. Sys.* **26**(1) (2006), 189–217.
[**15**] C. Gupta, W. Ott and A. Török. Memory loss for time-dependent piecewise expanding systems in higher dimension. *Math. Res. Lett.* **20**(1) (2013), 141–161.
[**16**] C. Kawan. Metric entropy of nonautonomous dynamical systems. *Nonauton. Dyn. Syst.* **1** (2014), 26–52.
[**17**] G. Keller and C. Liverani. Stability of the spectrum for transfer operators. *Ann. Sc. Norm. Super. Pisa Cl. Sci.* (4) **28**(1) (1999), 141–152.
[**18**] S. Kolyada, M. Misiurewicz and Ł. Snoha. Topological entropy of nonautonomous piecewise monotone dynamical systems on the interval. *Fund. Math.* **160**(2) (1999), 161–181.
[**19**] S. Kolyada and Ł. Snoha. Topological entropy of nonautonomous dynamical systems. *Random Comput. Dynam.* **4**(2–3) (1996), 205–233.
[**20**] A. Lasota and J. A. Yorke. When the long-time behavior is independent of the initial density. *SIAM J. Math. Anal.* **27**(1) (1996), 221–240.
[**21**] Y. Le Jan. On isotropic Brownian motions. *Z. Wahrscheinlichkeitstheor. Verwandte Geb.* **70**(4) (1985), 609–620.

https://doi.org/10.1017/etds.2016.9 Published online by Cambridge University Press

[**22**] V. P. Leonov. On the dispersion of time means of a stationary stochastic process. *Teor. Verojatnost. Primenen.* **6** (1961), 93–101.
[**23**] T. Lindvall. *Lectures on The Coupling Method*. Dover Publications, Inc., Mineola, NY, 2002, Corrected reprint of the 1992 original.
[**24**] A. N. Livšic. Certain properties of the homology of Y-systems. *Mat. Zametki* **10** (1971), 555–564.
[**25**] A. N. Livšic. Cohomology of dynamical systems. *Izv. Akad. Nauk SSSR Ser. Mat.* **36** (1972), 1296–1320.
[**26**] F. Mandl. *Statistical Physics (Manchester Physics Series)*, 2nd edn. John Wiley & Sons, 1988.
[**27**] A. Mohapatra and W. Ott. Memory loss for nonequilibrium open dynamical systems. *Discrete Contin. Dyn. Syst.* **34**(9) (2014), 3747–3759.
[**28**] P. Nándori, D. Szász and T. Varjú. A central limit theorem for time-dependent dynamical systems. *J. Stat. Phys.* **146**(6) (2012), 1213–1220.
[**29**] W. Ott, M. Stenlund and L.-S. Young. Memory loss for time-dependent dynamical systems. *Math. Res. Lett.* **16**(3) (2009), 463–475.
[**30**] L. C. G. Rogers and D. Williams. *Diffusions, Markov Processes, and Martingales. Vol. 2: Itô Calculus (Cambridge Mathematical Library)*. Cambridge University Press, Cambridge, 2000. Reprint of the second (1994) edition.
[**31**] M. Stenlund. Non-stationary compositions of Anosov diffeomorphisms. *Nonlinearity* **24** (2011), 2991–3018.
[**32**] M. Stenlund and H. Sulku. A coupling approach to random circle maps expanding on the average. *Stoch. Dyn.* **14**(4) (2014),1450008; 29 pp.
[**33**] M. Stenlund, L.-S. Young and H. Zhang. Dispersing billiards with moving scatterers. *Comm. Math. Phys.* **322**(3) (2013), 909–955.
[**34**] D. W. Stroock and S. R. Srinivasa Varadhan. *Multidimensional Diffusion Processes (Classics in Mathematics)*. Springer, Berlin, 2006. Reprint of the 1997 edition.
[**35**] L.-S. Young. Recurrence times and rates of mixing. *Israel J. Math.* **110** (1999), 153–188.
[**36**] Y. Zhu, Z. Liu, X. Xu and W. Zhang. Entropy of nonautonomous dynamical systems. *J. Korean Math. Soc.* **49**(1) (2012), 165–185.

https://doi.org/10.1017/etds.2016.9 Published online by Cambridge University Press