Probability Theory and Related Fields (2020) 178:735–770
https://doi.org/10.1007/s00440-020-00988-5

Superdiffusive limits for deterministic fast–slow dynamical systems

Ilya Chevyrev[1] · Peter K. Friz[2,3] · Alexey Korepanov[4] · Ian Melbourne[5]

Received: 21 August 2019 / Revised: 9 July 2020 / Published online: 16 July 2020

Abstract
We consider deterministic fast–slow dynamical systems on $\mathbb{R}^m \times Y$ of the form

$$\begin{cases} x_{k+1}^{(n)} = x_k^{(n)} + n^{-1}a\big(x_k^{(n)}\big) + n^{-1/\alpha}b\big(x_k^{(n)}\big)v(y_k), \\ y_{k+1} = f(y_k), \end{cases}$$

where $\alpha \in (1, 2)$. Under certain assumptions we prove convergence of the m-dimensional process $X_n(t) = x_{\lfloor nt \rfloor}^{(n)}$ to the solution of the stochastic differential equation

$$\mathrm{d}X = a(X)\mathrm{d}t + b(X) \diamond \mathrm{d}L_\alpha,$$

where L_α is an α-stable Lévy process and $\diamond$ indicates that the stochastic integral is in the Marcus sense. In addition, we show that our assumptions are satisfied for intermittent maps f of Pomeau–Manneville type.

Mathematics Subject Classification Primary 37A50; Secondary 60L20

✉ Ilya Chevyrev
ichevyrev@gmail.com

Peter K. Friz
friz@math.tu-berlin.de

Alexey Korepanov
a.korepanov@exeter.ac.uk

Ian Melbourne
i.melbourne@warwick.ac.uk

1 School of Mathematics, University of Edinburgh, Edinburgh EH9 3FD, UK

2 Institut für Mathematik, Technische Universität Berlin, Berlin, Germany

3 Weierstraß-Institut für Angewandte Analysis und Stochastik, Berlin, Germany

4 Department of Mathematics, University of Exeter, Exeter EX4 4QF, UK

5 Mathematics Institute, University of Warwick, Coventry CV4 7AL, UK

Contents

1 Introduction

Averaging and homogenisation for systems with multiple timescales is a longstanding and very active area of research [34]. We focus particularly on homogenisation, where the limiting equation is a stochastic differential equation (SDE). Recently there has been considerable interest in the case where the underlying multiscale system is deterministic, see [9–11,16,20,21,24,32,35] as well as our survey paper [8]. Almost all of this previous research has been concerned with the case where the limiting SDE is driven by Brownian motion. Here, we consider the case where the limiting SDE is driven by a superdiffusive α-stable Lévy process.

Let $\alpha \in (1, 2)$. The multiscale equations that we are interested in have the form

$$\begin{cases} x_{k+1}^{(n)} = x_k^{(n)} + n^{-1}a\big(x_k^{(n)}\big) + n^{-1/\alpha}b\big(x_k^{(n)}\big)v(y_k), \\ y_{k+1} = f(y_k) \end{cases} \tag{1.1}$$

defined on $\mathbb{R}^m \times Y$ where Y is a bounded metric space. Here

$$a\colon \mathbb{R}^m \to \mathbb{R}^m, \quad b\colon \mathbb{R}^m \to \mathbb{R}^{m\times d}, \quad v\colon Y \to \mathbb{R}^d, \quad f\colon Y \to Y.$$

It is assumed that the fast dynamical system $f\colon Y \to Y$ has an ergodic invariant probability measure μ and exhibits superdiffusive behaviour; specific examples for such f are described below. Let $v\colon Y \to \mathbb{R}^d$ be Hölder with $\int v\,d\mu = 0$. Define for $n \geq 1$,

$$W_n(t) = n^{-1/\alpha} \sum_{j=0}^{\lfloor nt \rfloor - 1} v \circ f^j. \tag{1.2}$$

Then W_n belongs to $D([0,1], \mathbb{R}^d)$, the Skorokhod space of càdlàg functions, and can be viewed as a random process on the probability space (Y, μ) depending on the initial condition $y_0 \in Y$. As $n \to \infty$, the sequence of random variables $W_n(1)$ converges weakly in $\mathbb{R}^d$ to an α-stable law, and the process W_n converges weakly in $D([0,1], \mathbb{R}^d)$ to the corresponding α-stable Lévy process L_α.

Now consider $x_0^{(n)} = \xi_n \in \mathbb{R}^m$, and solve (1.1) to obtain $(x_k^{(n)}, y_k)_{k \geq 0}$ depending on the initial condition $y_0 \in (Y, \mu)$. Define the càdlàg process $X_n \in D([0,1], \mathbb{R}^m)$ given by $X_n(t) = x_{\lfloor nt \rfloor}^{(n)}$; again we view this as a process on (Y, μ). Our aim is to show, under mild regularity assumptions on the functions $a \colon \mathbb{R}^m \to \mathbb{R}^m$ and $b \colon \mathbb{R}^m \to \mathbb{R}^{m \times d}$, that $X_n \to_w X$ where X is the solution of the SDE

$$\mathrm{d}X = a(X)\mathrm{d}t + b(X) \diamond \mathrm{d}L_\alpha, \qquad X(0) = \xi \tag{1.3}$$

and $\xi = \lim_{n \to \infty} \xi_n$. Here, $\diamond$ indicates that the SDE is in the Marcus sense [29] (see [2,5,25] for the general theory of Marcus SDEs and their applications).

Previously such a result was shown by Gottwald and Melbourne [16, Section 5] in the special case $d = m = 1$. Generally the method in [16] works provided the noise is exact, that is $d = m$ and $b = (Dr)^{-1}$ for some diffeomorphism $r \colon \mathbb{R}^m \to \mathbb{R}^m$, but cannot handle the general situation considered here where the noise term is typically not exact. There are three main complications:

(1) In the case of exact noise, it is possible to reduce to the case $b \equiv \mathrm{id}$ by a change of coordinates, similar to Wong–Zakai [45]. The general situation necessitates the use of alternative tools such as rough paths. In particular, weak convergence of W_n is no longer sufficient and we require in addition that W_n is tight in p-variation. This is shown in Theorem 1.3 below for specific examples, and in Sect. 6 for a large class of deterministic dynamical systems $f \colon Y \to Y$.

(2) Since the results for exact noise are achieved by a change of coordinates, the sense of convergence for W_n is inherited by X_n. However, in general, even if $W_n \to_w L_\alpha$ in one of the standard Skorokhod topologies [40], this need not be the case for X_n. This phenomenon already appears in the simplest situations, as illustrated in Example 1.4. Hence we have to consider convergence of X_n in generalised Skorokhod topologies as introduced recently in Chevyrev and Friz [7].

(3) Rigorous results on convergence to d-dimensional stable Lévy processes in deterministic dynamical systems are only available for $d = 1$, see [1,22,33,42]. Hence one of the aims of this paper is to extend the dynamical systems theory to cover the case $d \geq 2$. See Theorem 1.1 below for instances of this, and Sect. 6 for a general treatment.

In the remainder of the introduction, we discuss some of the issues associated to these three complications. We also mention some examples of fast dynamical systems that lead to superdiffusive behaviour. The archetypal such dynamical systems are the

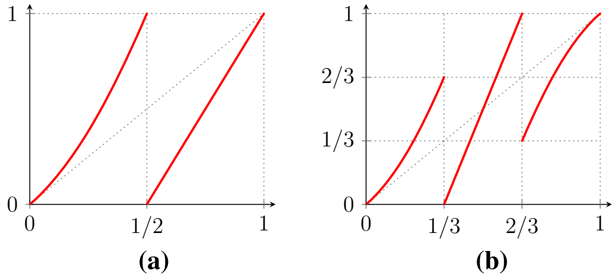

Fig. 1 Examples of intermittent maps: **a** the map (1.4), **b** the map (1.5)

intermittent maps introduced by Pomeau and Manneville [37]. Perhaps the simplest example [27] is the map $f: Y \to Y$, $Y = [0, 1]$, with a neutral fixed point at 0:

$$f(y) = \begin{cases} y(1 + 2^{1/\alpha} y^{1/\alpha}), & y \in [0, \frac{1}{2}), \\ 2y - 1, & y \in [\frac{1}{2}, 1]. \end{cases} \tag{1.4}$$

See Fig. 1a. Here, $\alpha > 0$ is a real parameter and there is a unique absolutely continuous invariant probability measure μ for $\alpha > 1$. Let $v: Y \to \mathbb{R}$ be Hölder with $\int_Y v \, d\mu = 0$ and $v(0) \neq 0$, and define W_n as in (1.2). For $\alpha \in (1, 2)$ it was shown by Gouëzel [17] (see also [46]) that $W_n(1)$ converges in distribution to an α-stable law. By Melbourne and Zweimüller [33], the process W_n converges weakly to the corresponding Lévy process L_α in the $\mathcal{M}_1$ Skorokhod topology on $D([0, 1], \mathbb{R})$.

Now let $d \geq 2$. There are two versions of the $\mathcal{M}_1$ topology on $D([0, 1], \mathbb{R}^d)$, see [43, Chapter 3.3]. In this paper we use the strong topology $\mathcal{SM}_1$. For $v: Y \to \mathbb{R}^d$ Hölder with $\int_Y v \, d\mu = 0$ and $v(0) \neq 0$, we prove convergence of W_n to a d-dimensional Lévy process L_α in the $\mathcal{SM}_1$ topology.

The example (1.4) is somewhat oversimplified for our purposes since L_α is essentially one-dimensional, being supported on the line $\{cv(0) : c \in \mathbb{R}\}$. This structure can be exploited in proving that $W_n \to_w L_\alpha$, though it is not clear if this simplifies the homogenisation result $X_n \to_w X$. To illustrate that we do not rely on one-dimensionality of the limiting process in any way, we consider an example with two neutral fixed points. (It is straightforward to extend to maps with a larger number of neutral fixed points.) Accordingly, our main example is the intermittent map $f: Y \to Y$, $Y = [0, 1]$, with two symmetric neutral fixed points at 0 and 1:

$$f(y) = \begin{cases} y\big(1 + 3^{1/\alpha} y^{1/\alpha}\big), & y \in [0, \frac{1}{3}), \\ 3y - 1, & y \in [\frac{1}{3}, \frac{2}{3}), \\ 1 - (1 - y)\big(1 + 3^{1/\alpha}(1 - y)^{1/\alpha}\big), & y \in [\frac{2}{3}, 1]. \end{cases} \tag{1.5}$$

See Fig. 1b. Again $\alpha > 0$ is a real parameter, there is a unique absolutely continuous invariant probability measure μ for $\alpha > 1$, and we restrict to the range $\alpha \in (1, 2)$.

As part of a result for a general class of nonuniformly expanding maps (Sect. 6) we prove:

Theorem 1.1 *Consider the intermittent map* (1.4) *or* (1.5) *with* $\alpha \in (1, 2)$ *and let* $v\colon Y \to \mathbb{R}^d$ *be Hölder with* $\int_Y v \, d\mu = 0$ *and* $v(0) \neq 0$*, also* $v(1) \neq 0$ *in case of* (1.5)*. Let* $\mathbb{P}$ *be any probability measure on* Y *that is absolutely continuous with respect to Lebesgue, and regard* W_n *as a process on* $(Y, \mathbb{P})$*. Then*

$$W_n \to_w L_\alpha \text{ in } D([0,1], \mathbb{R}^d) \text{ with the } \mathcal{SM}_1 \text{ topology as } n \to \infty,$$

where L_α *is a* d*-dimensional* α*-stable Lévy process.*

Remark 1.2 The limiting process L_α is explicitly identified in Sect. 6.2.

In the context of [16], the conclusion $W_n \to_w L_\alpha$ was sufficient to prove the homogenisation result $X_n \to_w X$. This is not the case for general noise, and we require tightness in p-variation. For $1 \le p < \infty$, recall that the *p-variation* of $u\colon [0,1] \to \mathbb{R}^d$ is given by

$$\|u\|_{p\text{-var}} = \sup_{0=t_0<t_1<\cdots<t_k=1} \left(\sum_{j=1}^{k} \big|u(t_j) - u(t_{j-1})\big|^p \right)^{1/p}, \tag{1.6}$$

where $|\cdot|$ denotes the Euclidean norm on $\mathbb{R}^d$.

Theorem 1.3 *Consider the intermittent map* (1.4) *or* (1.5) *with* $\alpha \in (1, 2)$ *and let* $v\colon Y \to \mathbb{R}^d$ *be Hölder with* $\int_Y v \, d\mu = 0$*. Let* $\mathbb{P}$ *be any probability measure on* Y *that is absolutely continuous with respect to Lebesgue. Then the family of random variables* $\|W_n\|_{p\text{-var}}$ *is tight on* $(Y, \mathbb{P})$ *for all* $p > \alpha$*.*

The main abstract result in this paper states that the properties established in Theorems 1.1 and 1.3 are the key ingredients required to solve the homogenisation problem. Informally:

> Consider the fast–slow system (1.1) and define W_n as in (1.2) and $X_n = x^{(n)}_{\lfloor nt \rfloor}$ with $x^{(n)}_0 = \xi_n$. Suppose that $\lim_{n\to\infty} \xi_n = \xi$, $W_n \to_w L_\alpha$, an α-stable Lévy process, in $D([0,1], \mathbb{R}^d)$ with the $\mathcal{SM}_1$-topology, and that $\|W_n\|_{p\text{-var}}$ is tight for all $p > \alpha$.
> If v is bounded and a, b are sufficiently smooth, then $X_n \to_w X$ in $D([0,1], \mathbb{R}^m)$ where X is the solution to the SDE (1.3).

We give a rigorous formulation of this result in Theorem 2.6 (in the above statement we assume that the limiting process is Lévy only for convenience—the result holds true for an arbitrary limiting process as seen from Theorem 2.6). To complete the statement, it is necessary to describe the topology on $D([0,1], \mathbb{R}^m)$ in which X_n converges. As already indicated, the $\mathcal{SM}_1$ topology is too strong in general. The next example illustrates where the problem lies.

Example 1.4 Let $\theta > 0$ and consider continuous deterministic processes $W_n\colon [0,1] \to \mathbb{R}$ which are equal to 0 on $[0, \frac{1}{2}]$, equal to θ on $[\frac{1}{2} + \frac{1}{n}, 1]$, and linear on $[\frac{1}{2}, \frac{1}{2} + \frac{1}{n}]$. Let $X_n = (X_n^1, X_n^2)$ be the solution of the ordinary differential equation

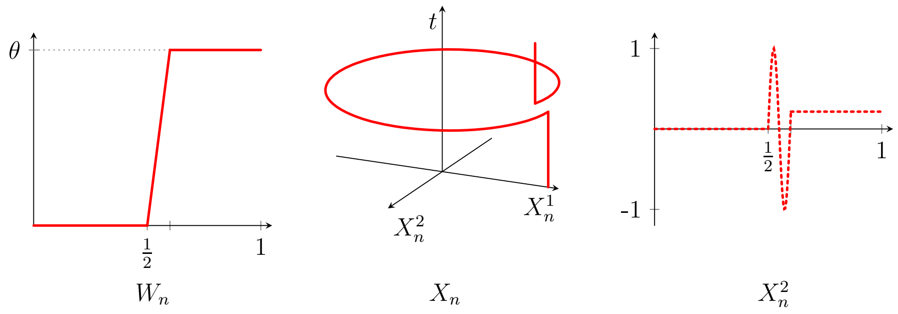

Fig. 2 Graphs of W_n and $X_n = (X_n^1, X_n^2)$ in Example 1.4

$$\begin{pmatrix} \mathrm{d}X_n^1 \\ \mathrm{d}X_n^2 \end{pmatrix} = \begin{pmatrix} -X_n^2 \\ X_n^1 \end{pmatrix} \mathrm{d}W_n, \qquad \begin{pmatrix} X_n^1(0) \\ X_n^2(0) \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

The graphs of W_n and X_n are shown in Fig. 2.

It is easy to see that W_n converges to $\theta\, 1_{[1/2,1]}$ in the $\mathcal{M}_1$ topology as $n \to \infty$, and that $(X_n^1, X_n^2) = (\cos W_n, \sin W_n)$. The process X_n converges pointwise to

$$X(t) = \begin{cases} (1,0), & t \leq \frac{1}{2}, \\ (\cos\theta, \sin\theta), & t > \frac{1}{2}. \end{cases}$$

In particular, if $\theta = 2\pi$, then $X \equiv (1, 0)$ is continuous. At the same time, X_n fails to converge in any of the Skorokhod topologies.

The problem outlined in Example 1.4 arises naturally in the fast–slow system (1.1). Figure 3 illustrates a realisation[1] of W_n and X_n for $d = m = 2$ and the map (1.5). The function b is taken as

$$b(x_1, x_2)\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix} v_1 + \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} v_2.$$

Note that, although W_n appears to converge in $\mathcal{SM}_1$ in accordance with Theorem 1.1, X_n moves along the integral curves of a vector field, and thus does not approximate its limit in $\mathcal{SM}_1$.

Topologies naturally suited for convergence in Example 1.4 were recently introduced in [7]. These topologies are a generalisation of the Skorokhod $\mathcal{SM}_1$ topology which allow for convenient control of differential equations. Briefly, jumps of a càdlàg process are interpreted as an instant travel along prescribed continuous paths which depend only on the start and end points of the jump. The full "pathspace" thus becomes the set of pairs (X, ϕ), where $X\colon [0,1] \to \mathbb{R}^d$ is a càdlàg path and ϕ is a so-called *path function* [6] which maps each jump $(X(t-), X(t))$ to a continuous path from $X(t-)$ to $X(t)$. It is often convenient to fix ϕ, which in turn determines a topology on

[1] Generated from https://khu.dedyn.io/work/scaled-graphs/fast-slow/.

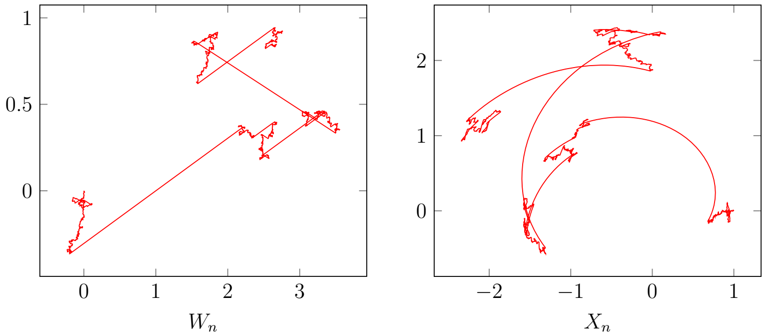

Fig. 3 Realisation of W_n and X_n with $n = 10^4$ points

càdlàg paths; if ϕ is linear, one recovers the $\mathcal{SM}_1$ topology. For our purposes, it is necessary to adapt the spaces introduced in [7], and we give details in Sects. 2 and 3.

The paper is organised as follows. In Sect. 2, we introduce the necessary prerequisites on generalised Skorokhod topologies and Marcus differential equations in order to state rigorously our main abstract result Theorem 2.6. The proof is given at the end of Sect. 3 after introducing the necessary results from rough path theory. In Sects. 4 to 6, we show that a class of nonuniformly expanding dynamical systems, including (1.4) and (1.5), satisfies the conclusions of Theorems 1.1 and 1.3 which are in turn the main hypotheses of Theorem 2.6. Section 4 deals with a class of uniformly expanding maps known as Gibbs–Markov maps, and Sect. 5 provides the inducing step to pass from uniformly expanding maps to nonuniformly expanding maps. In Sect. 6, we apply the results of Sects. 4 and 5 to the intermittent maps (1.4) and (1.5). The precise result on homogenisation of the system (1.1) with fast dynamics given by either (1.4) or (1.5) is stated in Corollary 6.4.

Notation We use "big O" and $\lesssim$ notation interchangeably, writing $a_n = O(b_n)$ or $a_n \lesssim b_n$ if there is a constant $C > 0$ such that $a_n \le Cb_n$ for all sufficiently large n. As usual, $a_n = o(b_n)$ means that $\lim_{n\to\infty} a_n/b_n = 0$ and $a_n \sim b_n$ means that $\lim_{n\to\infty} a_n/b_n = 1$.

2 Setup and result

In this section, we collect the material necessary to formulate our main abstract result Theorem 2.6.

2.1 Skorokhod topologies

Let $D = D([0, 1], \mathbb{R}^d)$ denote the Skorokhod space of càdlàg functions, i.e. the set of functions $X : [0, 1] \to \mathbb{R}^d$ which are right-continuous with left limits. For $X \in D$ and $t \in [0, 1]$, we denote $X(t-) = \lim_{s\nearrow t} X(s)$, with the convention that $X(0-) = X(0)$.

Let Λ denote the set of all increasing bijections $\lambda\colon [0,1] \to [0,1]$ and let $\mathrm{id} \in \Lambda$ denote the identity map $\mathrm{id}(t) = t$. For $X_1, X_2 \in D$, let $\boldsymbol{\sigma}_\infty(X_1, X_2)$ be the Skorokhod distance

$$\boldsymbol{\sigma}_\infty(X_1, X_2) = \inf_{\lambda \in \Lambda} \max\{\|\lambda - \mathrm{id}\|_\infty, \|X_1 \circ \lambda - X_2\|_\infty\},$$

where $\|X\|_\infty = \sup_{t\in[0,1]} |X(t)|$. The topology on D induced by $\boldsymbol{\sigma}_\infty$ is known as the strong $\mathcal{J}_1$, or $\mathcal{S}\mathcal{J}_1$, topology.

Another important topology on D is the strong $\mathcal{M}_1$, or $\mathcal{S}\mathcal{M}_1$, topology defined as follows. For $X \in D$ consider the "completed" graph $\Gamma(X) = \{(t, x) \in [0,1] \times \mathbb{R}^d : x \in [X(t-), X(t)]\}$, and let $\Lambda^*(X)$ be the set of all continuous bijections $(\lambda, \gamma)\colon [0,1] \to \Gamma(X)$ with $\lambda(0) = 0$. Then the $\mathcal{S}\mathcal{M}_1$ topology on D is induced by the metric

$$d_{\mathcal{S}\mathcal{M}_1}(X_1, X_2) = \inf_{\substack{(\lambda_i,\gamma_i)\in\Lambda^*(X_i)\\ i=1,2}} \max\{\|\lambda_1 - \lambda_2\|_\infty, \|\gamma_1 - \gamma_2\|_\infty\}.$$

2.2 Generalised $\mathcal{S}\mathcal{M}_1$ topologies

We now introduce generalisations of the $\mathcal{S}\mathcal{M}_1$ topology from [7].

For $1 \leq p < \infty$, recall the p-variation $\|u\|_{p\text{-var}}$ of $u\colon [0,1] \to \mathbb{R}^d$ defined by (1.6). We furthermore denote $|||u|||_{p\text{-var}} = |u(0)| + \|u\|_{p\text{-var}}$. Let

$$D^{p\text{-var}} = \big\{u \in D\big([0,1], \mathbb{R}^d\big) : \|u\|_{p\text{-var}} < \infty\big\}$$

and $C^{p\text{-var}}([0,1], \mathbb{R}^d) \subset D^{p\text{-var}}$ be the set of $u \in D^{p\text{-var}}$ which are continuous. Let $\boldsymbol{\sigma}_{p\text{-var}}$ denote the Skorokhod-type p-variation on $D^{p\text{-var}}$:

$$\boldsymbol{\sigma}_{p\text{-var}}(X_1, X_2) = \inf_{\lambda\in\Lambda} \max\{\|\lambda - \mathrm{id}\|_\infty, |||X_1 \circ \lambda - X_2|||_{p\text{-var}}\}.$$

Definition 2.1 A *path function* on $\mathbb{R}^d$ is a map $\phi\colon J \to C([0,1], \mathbb{R}^d)$, where $J \subset \mathbb{R}^d \times \mathbb{R}^d$, for which $\phi(x, y)(0) = x$ and $\phi(x, y)(1) = y$ for all $(x, y) \in J$. For a path $X \in D([0,1], \mathbb{R}^d)$, we say that $t \in [0,1]$ is a jump time of X if $X(t-) \neq X(t)$. A pair (X, ϕ) is called admissible if all the jumps of X are in the domain of definition of ϕ, i.e. $(X(t-), X(t)) \in J$ for all jump times t of X. We denote by $\bar{\mathscr{D}}([0,1], \mathbb{R}^d)$ the space of admissible pairs (X, ϕ). We let $\mathscr{D}([0,1], \mathbb{R}^d) = \bar{\mathscr{D}}([0,1], \mathbb{R}^d)/\sim$, where $(X_1, \phi_1) \sim (X_2, \phi_2)$ if $X_1 = X_2$ and $\phi_1(X_1(t-), X_1(t))$ is a reparametrisation of $\phi_2(X_1(t-), X_1(t))$ for all jump times t of X_1.

Remark 2.2 We often keep implicit the interval $[0,1]$ and $\mathbb{R}^d$, as well as J, when they are clear from the context. We allow J to be a strict subset of $\mathbb{R}^d \times \mathbb{R}^d$ since this case arises naturally when considering driver–solution pairs for canonical differential equations, see the final discussion in Sect. 2.3.

A simple path function which shall play an important role is the following.

Definition 2.3 The *linear path function* on $\mathbb{R}^k$ is the map $\ell_k : \mathbb{R}^k \times \mathbb{R}^k \to C([0,1],\mathbb{R}^k)$ defined by $\ell_k(x,y)(t) = x + t(y-x)$ for all $x, y \in \mathbb{R}^k$.

Fix a sequence $r_1, r_2, \ldots > 0$ with $\sum_j r_j < \infty$. Given $(X,\phi) \in \bar{\mathscr{D}}$ and $\delta > 0$, let $X^{\phi,\delta} \in C([0,1],\mathbb{R}^d)$ denote the continuous version of X, where the k-th largest jump is made continuous using ϕ on a fictitious time interval of length δr_k. More precisely:

- Let $m \ge 0$ be the number of jumps (possibly infinite) of X. We order the jump times $\{t_j\}_{j=1}^m$ so that $|X(t_k) - X(t_k-)| \ge |X(t_{k+1}) - X(t_{k+1}-)|$ for each k, with $t_k < t_{k+1}$ in case of equality.
- Let $r = \sum_{j=1}^m r_j$ and define the map

$$\tau \colon [0,1] \to [0, 1+\delta r], \quad \tau(t) = t + \sum_k \delta r_k 1_{\{t_k \le t\}}. \tag{2.1}$$

- Define an intermediate process $\hat{X} \in C([0,1+\delta r],\mathbb{R}^d)$,

$$\hat{X}(t) = \begin{cases} X(s) & \text{if } t = \tau(s) \text{ for some } s \in [0,1], \\ \phi(X(t_k-), X(t_k))\big(\frac{s-\tau(t_k-)}{\delta r_k}\big) & \text{if } t \in [\tau(t_k-), \tau(t_k)) \text{ for some } k. \end{cases}$$

- Finally, let $X^{\phi,\delta}(t) = \hat{X}(t(1+\delta r))$, scaling the domain of $\hat{X}$ from $[0,1+\delta r]$ to $[0,1]$.

For $(X,\phi) \in \mathscr{D}([0,1],\mathbb{R}^d)$ and $p \ge 1$, let

$$\|(X,\phi)\|_{p\text{-var}} = \|X^{\phi,1}\|_{p\text{-var}}.$$

Note that $\|(X,\phi)\|_{p\text{-var}}$ is well-defined since $\|X^{\phi,1}\|_{p\text{-var}}$ depends on neither the parametrisation of ϕ, nor the sequence $\{r_k\}$. Let

$$\mathscr{D}^{p\text{-var}} = \{(X,\phi) \in \mathscr{D} : \|(X,\phi)\|_{p\text{-var}} < \infty\}.$$

Given (X_1,ϕ_1) and (X_2,ϕ_2) in $\mathscr{D}^{p\text{-var}}$, let

$$\boldsymbol{\alpha}_{p\text{-var}}((X_1,\phi_1),(X_2,\phi_2)) = \lim_{\delta\to 0} \boldsymbol{\sigma}_{p\text{-var}}\big(X_1^{\phi_1,\delta}, X_2^{\phi_2,\delta}\big),$$

which defines a metric on $\mathscr{D}^{p\text{-var}}$ [7, Remark 3.8].

2.3 Marcus differential equations

For $\gamma > 0$, let $C^\gamma(\mathbb{R}^m,\mathbb{R}^n)$ denote the space of functions $b \colon \mathbb{R}^m \to \mathbb{R}^n$ such that

$$\|b\|_{C^\gamma} = \max_{|\alpha|=0,\ldots,\lfloor\gamma\rfloor} \|D^\alpha b\|_\infty + \sup_{x,y\in\mathbb{R}^m} \max_{|\alpha|=\lfloor\gamma\rfloor} \frac{|D^\alpha b(x) - D^\alpha b(y)|}{|x-y|^{\gamma-\lfloor\gamma\rfloor}} < \infty.$$

Note that our notation is slightly non-standard since $b \in C^N$ for $N \in \mathbb{N}$ implies only that the $(N-1)$-th derivative of b is Lipschitz rather than continuous.

Suppose that $W \in D^{p\text{-var}}([0,1], \mathbb{R}^d)$ with $1 \le p < 2$, and that $a \in C^\beta(\mathbb{R}^m, \mathbb{R}^m)$ and $b \in C^\gamma(\mathbb{R}^m, \mathbb{R}^{m\times d})$ with $\beta > 1$ and $\gamma > p$. Under these conditions, we can define and solve (in a purely deterministic way) a Marcus-type differential equation

$$\mathrm{d}X = a(X)\mathrm{d}t + b(X) \diamond \mathrm{d}W. \tag{2.2}$$

The solution is obtained as follows from the theory of continuous rough differential equations (RDEs) in the Young regime [12,14,28]. Consider the càdlàg path $\widetilde{W} \colon [0,1] \to \mathbb{R}^{1+d}$ given by $\widetilde{W}(t) = (t, W(t))$. Using the notation of Sect. 2.2, consider the continuous path $\widetilde{W}^{\phi,1} \colon [0, 1+r] \to \mathbb{R}^d$, where $\phi = \ell_{1+d}$ is the linear path function on $\mathbb{R}^{1+d}$. Let $\tau \colon [0,1] \to [0, 1+r]$ be the corresponding map given by (2.1). Then $\|\widetilde{W}^{\phi,1}\|_{p\text{-var}} \lesssim \|W\|_{p\text{-var}}$ (see e.g. [6, Corollary A.6]), and therefore one can solve the (continuous) RDE

$$\mathrm{d}\widetilde{X} = (a, b)(\widetilde{X})\mathrm{d}\widetilde{W}.$$

The solution is a continuous path $\widetilde{X} \colon [0, 1+r] \to \mathbb{R}^m$ of finite p-variation. The solution to (2.2) is the càdlàg path $X \colon [0,1] \to \mathbb{R}^d$ given by $X(t) = \widetilde{X}(\tau(t))$. We discuss a more general interpretation of this equation in Sect. 3.2.

Remark 2.4 In the case that W is a semimartingale, one can verify that X is the solution to the classical Marcus SDE (see [7, Proposition 4.16] for the general case $p > 2$ but with stronger regularity assumptions on a, b; the proof carries over to our setting without change).

To properly describe solutions of (2.2) and regularity of the solution map $W \mapsto X$, it is not enough to look at X as an element of $D([0,1], \mathbb{R}^m)$. As in Example 1.4, one may have $X \equiv 0$ say, but with sizeable jumps in fictitious time.

Following [7], we consider the driver-solution space $D([0,1], \mathbb{R}^{d+m})$, made to contain the pairs (W, X), and introduce a new path function on $\mathbb{R}^{d+m}$.

Definition 2.5 Consider $b \in C^1(\mathbb{R}^m, \mathbb{R}^{m\times d})$. For $x \in \mathbb{R}^m$ and $\Phi \in C^{1\text{-var}}([0,1], \mathbb{R}^d)$, let $\pi_b[x; \Phi] \in C^{1\text{-var}}([0,1], \mathbb{R}^m)$ denote the solution Π of the equation

$$\mathrm{d}\Pi = b(\Pi)\mathrm{d}\Phi, \quad \Pi(0) = x.$$

We define the path function ϕ_b on $\mathbb{R}^{d+m}$ by

$$\phi_b\big((w_1, x_1), (w_2, x_2)\big)(t) = \big(\ell_d(w_1, w_2)(t), \pi_b[x_1; \ell_d(w_1, w_2)](t)\big), \tag{2.3}$$

which is defined on

$$J_b = \big\{\big((w_1, x_1), (w_2, x_2)\big) : w_1, w_2 \in \mathbb{R}^d,\ \pi_b[x_1; \ell_d(w_1, w_2)](1) = x_2\big\}.$$

Note that J_b is a strict subset of $\mathbb{R}^{d+m} \times \mathbb{R}^{d+m}$. Observe that if X solves (2.2), then $((W, X), \phi_b) \in \mathscr{D}^{p\text{-var}}([0, 1], \mathbb{R}^{d+m})$ and the path function ϕ_b describes how the discontinuities of (W, X) are traversed in fictitious time.

2.4 Main abstract result

Now we are ready for a rigorous formulation of the main abstract result. Consider the fast–slow system (1.1) with initial condition $x_0^{(n)} = \xi_n$ such that $\lim_{n\to\infty} \xi_n = \xi$. Suppose that $\alpha \in (1, 2)$, $\alpha' \in [\alpha, 2)$, $v \in L^\infty(Y, \mathbb{R}^d)$, $a \in C^\beta(\mathbb{R}^m, \mathbb{R}^m)$, $b \in C^\gamma(\mathbb{R}^m, \mathbb{R}^{m\times d})$ for some $\beta > 1$, $\gamma > \alpha'$. Define W_n as in (1.2) and $X_n(t) = x^{(n)}_{\lfloor nt\rfloor}$.

Theorem 2.6 *Suppose that*

- $W_n \to_w L$ *in* $D([0, 1], \mathbb{R}^d)$ *with the* $\mathcal{SM}_1$ *topology as* $n \to \infty$ *for some process* L.
- $\|W_n\|_{p\text{-var}}$ *is tight for all* $p > \alpha'$.

Then, for all $p > \alpha'$, *it holds that* $\|L\|_{p\text{-var}} < \infty$ *a.s. and*

$$((W_n, X_n), \ell_{d+m}) \to_w ((L, X), \phi_b) \quad \text{as} \quad n \to \infty$$

in $(\mathscr{D}^{p\text{-var}}([0, 1], \mathbb{R}^{d+m}), \boldsymbol{\alpha}_{p\text{-var}})$, *where* X *is the solution of the Marcus differential equation*

$$\mathrm{d}X = a(X)\mathrm{d}t + b(X) \diamond \mathrm{d}L, \qquad X(0) = \xi \in \mathbb{R}^m. \tag{2.4}$$

The proof of Theorem 2.6 is given at the end of Sect. 3.

Remark 2.7 (a) The property $\|L\|_{p\text{-var}} < \infty$ a.s. together with $\gamma > \alpha'$ guarantees that the Marcus equation (2.4) admits a unique solution for a.e. realisation of L. In our applications, L is an α-stable Lévy process, for which the finiteness of $\|L\|_{p\text{-var}}$ is classical, and we take $\alpha' = \alpha$. We introduce the parameter α' to highlight that the threshold for the value of p in the second condition of Theorem 2.6 does not need to be the same α as in (1.2).

(b) The drift vector field a plays no role in the definition of ϕ_b. This is expected since the driver $V_n(t) = n^{-1}\lfloor tn\rfloor$ corresponding to a in the RDE solved by X_n (see the proof of Theorem 2.6 below) converges in q-variation for every $q > 1$ to a process with no jumps.

(c) Since the limiting process L in general has jumps, it is crucial that we pair (L, X) with the path function ϕ_b. In contrast, the jumps of (W_n, X_n) are of magnitude at most $n^{-1/\alpha}$, so (W_n, X_n) is almost a continuous path for large n; we make the reference to ℓ_{d+m} only for convenience (cf. (3.10) below).

Recall that a stochastic process $(L_t)_{t\in[0,1]}$ is called *stochastically continuous* if, for all $t \in [0, 1]$, $L_s \to L_t$ in probability as $s \to t$. Note that Lévy processes are stochastically continuous by definition.

Corollary 2.8 *In the setting of Theorem 2.6, suppose further that the process* L *is stochastically continuous. Then* $X_n \to X$ *in the sense of finite dimensional distributions.*

Proof Consider $0 \leq t_1 < \cdots < t_k \leq 1$. The map

$$(Y, \phi) \mapsto (Y(t_1), \ldots, Y(t_k)), \qquad \left(\mathscr{D}^{p\text{-var}}([0,1], \mathbb{R}^{d+m}), \boldsymbol{\alpha}_{p\text{-var}}\right) \to \mathbb{R}^{(d+m)k} \quad (2.5)$$

is continuous at (Y, ϕ) whenever the path Y is continuous at all t_j, see [7, Lemma 2.12]. Furthermore, if $t \in [0,1]$ is a continuity point of L, then it is also a continuity point of the solution X to (2.4). Since L is càdlàg and stochastically continuous, any fixed $t \in [0,1]$ is a.s. a continuity point of L (see e.g. the proof of [2, Lemma 2.3.2]), $((L, X), \phi_b)$ is a.s. a continuity point of the map (2.5). In particular, by Theorem 2.6 and the continuous mapping theorem, $(X_n(t_1), \ldots, X_n(t_k))$ converges in law to $(X(t_1), \ldots, X(t_k))$, as required. □

Remark 2.9 As in Example 1.4, we do not expect that $X_n \to_w X$ in any of the Skorokhod topologies, or that $f(X_n) \to_w f(X)$ for certain standard functionals $f\colon D \to \mathbb{R}$ that are continuous with respect to the Skorokhod topologies, such as $f(X) = \|X\|_\infty$. Instead we have for example that $\|\widetilde{X}_n\|_\infty \to_w \|\widetilde{X}\|_\infty$, where $\widetilde{X}_n$ and $\widetilde{X}$ are the corresponding components of the continuous paths $(W_n, X_n)^{\ell_{d+m},1}$ and $(W, X)^{\phi_b,1}$.

3 Rough path formulation

In this section we expand the material in Sect. 2 in order to formulate and prove an abstract convergence result, Theorem 3.4, from which Theorem 2.6 follows.

3.1 Generalised $\mathcal{SM}_1$ topologies with mixed variation

We use a modified version of the topologies from [7] suitable for handling differential equations with drift. We continue using notation from Sect. 2.

For $1 \leq p, q < \infty$, we define the mixed (q, p)-variation for $u = (u^0, u^1, \ldots, u^d) = (u^0, \bar{u})\colon [0,1] \to \mathbb{R}^{1+d}$ by

$$\|u\|_{(q,p)\text{-var}} = \|u^0\|_{q\text{-var}} + \|\bar{u}\|_{p\text{-var}}.$$

Let

$$D^{(q,p)\text{-var}} = \left\{u \in D\big([0,1], \mathbb{R}^{1+d}\big) : \|u\|_{(q,p)\text{-var}} < \infty\right\}$$

and $C^{(q,p)\text{-var}}([0,1], \mathbb{R}^{1+d}) \subset D^{(q,p)\text{-var}}$ be the set of $u \in D^{(q,p)\text{-var}}$ which are continuous. We furthermore denote $|||u|||_{(q,p)\text{-var}} = |u(0)| + \|u\|_{(q,p)\text{-var}}$ and define

$$\boldsymbol{\sigma}_{(q,p)\text{-var}}(X_1, X_2) = \inf_{\lambda \in \Lambda} \max\{\|\lambda - \mathrm{id}\|_\infty, |||X_1 \circ \lambda - X_2|||_{(q,p)\text{-var}}\}.$$

Given (X_1, ϕ_1) and (X_2, ϕ_2) in $\bar{\mathscr{D}}$, let

$$\boldsymbol{\alpha}_\infty((X_1, \phi_1), (X_2, \phi_2)) = \lim_{\delta \to 0} \boldsymbol{\sigma}_\infty\big(X_1^{\phi_1,\delta}, X_2^{\phi_2,\delta}\big).$$

Following [7, Lemma 2.7], the limit exists, is independent of the choice of the sequence r_k, and is invariant under reparametrisation of the path functions. In particular, $\boldsymbol{\alpha}_\infty$ induces a pseudometric on $\mathscr{D}$.

For $(X, \phi) \in \mathscr{D}([0,1], \mathbb{R}^{1+d})$, let

$$\|(X, \phi)\|_{(q,p)\text{-var}} = \|X^{\phi,1}\|_{(q,p)\text{-var}}.$$

As before, note that $\|(X, \phi)\|_{(q,p)\text{-var}}$ is well-defined since $\|X^{\phi,1}\|_{(q,p)\text{-var}}$ does not depend on the parametrisation of ϕ, nor the sequence $\{r_k\}$. Let

$$\mathscr{D}^{(q,p)\text{-var}} = \{(X, \phi) \in \mathscr{D} : \|(X, \phi)\|_{(q,p)\text{-var}} < \infty\}.$$

Given (X_1, ϕ_1) and (X_2, ϕ_2) in $\mathscr{D}^{(q,p)\text{-var}}$, let

$$\boldsymbol{\alpha}_{(q,p)\text{-var}}((X_1, \phi_1), (X_2, \phi_2)) = \lim_{\delta \to 0} \boldsymbol{\sigma}_{(q,p)\text{-var}}\big(X_1^{\phi_1,\delta}, X_2^{\phi_2,\delta}\big),$$

which is well-defined and induces a metric on $\mathscr{D}^{(q,p)\text{-var}}$ (cf. [7, Remark 3.8]).

3.2 Differential equations with càdlàg drivers

For $\beta, \gamma > 0$, denote by $C^{\beta,\gamma}$ the space of all $b = (b^0, b^1, \dots, b^d) \colon \mathbb{R}^m \to \mathbb{R}^{m\times(1+d)}$ such that

$$\|b\|_{C^{\beta,\gamma}} = \|b^0\|_{C^\beta} + \max_{i=1,\dots,d} \|b^i\|_{C^\gamma} < \infty.$$

Suppose $1 \le q \le p < 2$ and that $b \in C^{\beta,\gamma}$ with $\beta > q$ and $\gamma > p$ such that

$$\frac{\beta - 1}{p} + \frac{1}{q} > 1 \quad \text{and} \quad \frac{\gamma - 1}{q} + \frac{1}{p} > 1. \tag{3.1}$$

Remark 3.1 See [14, Remark 12.7] for a discussion about condition (3.1). In our applications, we will consider $\beta > 1$ and $\gamma > p$ as fixed, and $q = 1+\kappa$ for $\kappa > 0$ arbitrarily small. In this case condition (3.1) is always attained by taking κ sufficiently small, which explains why it does not appear in Theorem 2.6.

Recall that under these conditions, if $W \in C^{(q,p)\text{-var}}([0,1], \mathbb{R}^{1+d})$, then the canonical RDE (in the Young regime)

$$\mathrm{d}X = b(X)\mathrm{d}W, \qquad X(0) = \xi \in \mathbb{R}^m$$

admits a unique solution $X \in C^{p\text{-var}}([0,1], \mathbb{R}^m)$.

For general $W \in D^{(q,p)\text{-var}}([0,1], \mathbb{R}^{1+d})$, consider the RDE

$$\mathrm{d}X = b(X) * \mathrm{d}W, \qquad X(0) = \xi. \tag{3.2}$$

Here, $*$ stands for one of the different ways to interpret a differential equation in the presence of discontinuities, which in general result in different solutions X. Two common choices (considered in the case $q = p$ by Williams [44] and studied further in [6,7,13,15]) are

- *Geometric (Marcus) RDE.* The solution is completely analogous to that of (2.2): we solve the continuous RDE $\mathrm{d}\widetilde{X} = b(\widetilde{X})\mathrm{d}W^{\phi,1}$, where $\phi = \ell_{1+d}$ is the linear path function on $\mathbb{R}^{1+d}$, and then remove the fictitious time intervals (note that the RDE is well-posed since $\|W^{\phi,1}\|_{(q,p)\text{-var}} \lesssim \|W\|_{(q,p)\text{-var}}$ by Chevyrev [6, Corollary A.6]). For geometric RDEs we use the notation
$$\mathrm{d}X = b(X) \diamond \mathrm{d}W, \qquad X(0) = \xi. \tag{3.3}$$
Observe that $((W, X), \phi_b) \in \mathscr{D}^{(q,p)\text{-var}}([0, 1], \mathbb{R}^{1+d+m})$, where ϕ_b is the path function on $\mathbb{R}^{1+d+m}$ as in Definition 2.5 with ℓ_d replaced by ℓ_{1+d}.
- *Forward (Itô) RDE.* The solution satisfies the integral equation
$$X(t) = X(0) + \int_0^t b(X(s-))\mathrm{d}W(s), \tag{3.4}$$
where the integral is understood as a limit of Riemann–Stieltjes sums with $b(X(s-))$ evaluated at the left limit points of the partition intervals:
$$\int_0^t b(X(s-))\mathrm{d}W(s) = \lim_{|\mathcal{P}|\to 0} \sum_{[s,s']\in\mathcal{P}} b(X(s-))(W(s') - W(s)).$$
Here, $\mathcal{P}$ are partitions of $[0, t]$ into intervals, and $|\mathcal{P}|$ is the size of the longest interval. For forward RDEs we use the notation
$$\mathrm{d}X = b(X)^{-}\mathrm{d}W, \qquad X(0) = \xi.$$

Remark 3.2 Geometric RDEs use linear paths to connect the endpoints of each jump. As mentioned in the introduction, this has been generalised in [7] allowing one to solve
$$\mathrm{d}X = b(X) \diamond \mathrm{d}(W, \phi), \qquad X(0) = \xi, \tag{3.5}$$
for any $(W, \phi) \in \mathscr{D}^{(q,p)\text{-var}}([0, 1], \mathbb{R}^{1+d})$. The interpretation is as for geometric RDEs: we construct a continuous path, solve the canonical RDE $\mathrm{d}\widetilde{X} = b(\widetilde{X})\mathrm{d}W^{\phi,1}$, and then remove fictitious time intervals. Then $((W, X), \phi_b) \in \mathscr{D}^{(q,p)\text{-var}}([0, 1], \mathbb{R}^{1+d+m})$, where ϕ_b is the path function on $\mathbb{R}^{1+d+m}$ as in Definition 2.5 with ℓ_d replaced by ϕ, and the solution map of (3.5)
$$\mathbb{R}^m \times \left(\mathscr{D}^{(q,p)\text{-var}}([0, 1], \mathbb{R}^{1+d}), \boldsymbol{\alpha}_{(q,p)\text{-var}}\right) \to \left(\mathscr{D}^{(q,p)\text{-var}}([0, 1], \mathbb{R}^{1+d+m}), \boldsymbol{\alpha}_{(q,p)\text{-var}}\right),$$
$$(\xi, (W, \phi)) \mapsto ((W, X), \phi_b)$$
is locally Lipschitz continuous. (These results were shown in [7, Theorem 3.13] for $q = p$, but the same proof applies mutatis mutandis for the general case upon using the

RDE with drift estimates [14, Theorem 12.10]. In fact one can allow rough path drivers in $\mathbb{R}^{d'+d}$ with finite (q, p)-variation for arbitrary $p, q \geq 1$ satisfying $p^{-1} + q^{-1} > 1$. We consider only $d' = 1$ and $1 \leq q \leq p < 2$ since this suffices for our purposes.)

3.3 Convergence of forward RDEs to geometric RDEs

For the remainder of this section, let us fix $1 \leq q \leq p < 2$, $\beta > q$, $\gamma > p$, such that (3.1) holds. Suppose that $W \in D^{(q,p)\text{-var}}([0,1], \mathbb{R}^{1+d})$ and $b \in C^{\beta,\gamma}$. Then for every $\xi \in \mathbb{R}^m$, the geometric RDE

$$\mathrm{d}\widetilde{X} = b(\widetilde{X}) \diamond \mathrm{d}W, \qquad \widetilde{X}(0) = \xi$$

admits a unique solution $\widetilde{X} \in D^{p\text{-var}}([0,1], \mathbb{R}^m)$.

Suppose now that W has finitely many jumps at times $0 < t_1 < \cdots < t_n \leq 1$. Then the solution X of the forward RDE

$$\mathrm{d}X = b(X)^{-}\mathrm{d}W, \qquad X(0) = \xi$$

can be obtained by solving the canonical RDE on each of the intervals $[0, t_1)$, $[t_1, t_2)$, $\dots [t_n, 1)$ (on which W is continuous), and requiring that at the jump times

$$X(t_k) = X(t_k-) + b(X(t_k-))(W(t_k) - W(t_k-)). \tag{3.6}$$

Hence in the case that W has finitely many jumps, it is straightforward to construct the solution X first on $[0, t_1)$, then at t_1, then on $[t_1, t_2)$ and so on. As we shall see, this construction furthermore allows for an easy extension of stability results of continuous RDEs to the setting with jumps.

Remark 3.3 The construction of the forward solution for processes with infinitely many discontinuities is more involved, and can be achieved by solving directly the integral equation (3.4). This is done in [15] but is not required here.

Recall that ϕ_b is the path function on $\mathbb{R}^{1+d+m}$ as in Definition 2.5 with ℓ_d replaced by ℓ_{1+d}.

Theorem 3.4 *Suppose that $\{W_n\}_{n\geq 1}$ is a sequence of $D^{(q,p)\text{-var}}([0,1], \mathbb{R}^{1+d})$-valued random elements with almost surely finitely many jumps. Suppose that $b \in C^{\beta,\gamma}$. Let X_n be the solution of the forward RDE*

$$\mathrm{d}X_n = b(X_n)^{-}\mathrm{d}W_n, \qquad X_n(0) = \xi_n \in \mathbb{R}^m.$$

Suppose that

- (a) $\lim_{n\to\infty} \xi_n = \xi$ *for some* $\xi \in \mathbb{R}^m$,
- (b) $W_n \to_w W$ *in* $D([0,1], \mathbb{R}^{1+d})$ *with the* $\mathcal{SM}_1$ *topology as* $n \to \infty$ *(we allow the limit process* W *to have infinitely many jumps),*
- (c) *the family of random variables* $\|W_n\|_{(q,p)\text{-var}}$ *is tight,*

(d) $\sum_t |W_n(t) - W_n(t-)|^2 \to_w 0$ *as* $n \to \infty$, *where the sum is over all jump times of* W_n.

Then $\|W\|_{(q,p)\text{-var}} < \infty$ *almost surely. Let* X *be the solution of the geometric RDE*

$$\mathrm{d}X = b(X) \diamond \mathrm{d}W, \qquad X(0) = \xi.$$

(The RDE is well-posed because $\|W\|_{(q,p)\text{-var}} < \infty$*.) Then for each* $q' > q$ *and* $p' > p$,

$$((W_n, X_n), \ell_{1+d+m}) \to_w ((W, X), \phi_b) \quad \text{in} \quad \left(\mathscr{D}([0,1], \mathbb{R}^{1+d+m}), \boldsymbol{\alpha}_{(q',p')\text{-var}}\right)$$

as $n \to \infty$.

We give the proof after several preliminary results. We will see that if X_n solved the geometric RDE $\mathrm{d}X_n = b(X_n) \diamond \mathrm{d}W_n$ instead of the forward RDE, then Theorem 3.4 would readily follow from [7] (and assumption (d) would not be needed). In Lemma 3.6, we verify that under assumption (d) the solution of the forward RDE $\mathrm{d}X_n = b(X_n)^- \mathrm{d}W_n$ closely approximates the solution of the geometric RDE $\mathrm{d}X_n = b(X_n) \diamond \mathrm{d}W_n$ (generalising a result of [44]). First we show how a single jump of a geometric solution relates to a "forward" jump (cf. [44, Lemma 1.1, Eq. (11)]). Define the semi-norm

$$\|b\|_{\mathrm{Lip}} = \sup_{x,y \in \mathbb{R}^m} \frac{|b(x) - b(y)|}{|x - y|}.$$

Lemma 3.5 *Suppose that* $X \in C([0,1], \mathbb{R}^m)$ *solves the ODE* $\mathrm{d}X = b(X)\mathrm{d}t$ *with* b *Lipschitz. Then* $\left|X(1) - X(0) - b(X(0))\right| \le \|b\|_{\mathrm{Lip}} \|b\|_\infty / 2$.

Proof Write $X(1) = X(0) + b(X(0)) + \int_0^1 \big(b(X(t)) - b(X(0))\big)\mathrm{d}t$. Since $|X(t) - X(0)| \le \|b\|_\infty t$,

$$\left|\int_0^1 \big(b(X(t)) - b(X(0))\big)\mathrm{d}t\right| \le \|b\|_{\mathrm{Lip}} \int_0^1 |X(t) - X(0)|\mathrm{d}t \le \|b\|_{\mathrm{Lip}} \|b\|_\infty \int_0^1 t\mathrm{d}t.$$

□

We now quantify the error in moving from forward to geometric solutions.

Lemma 3.6 *Suppose that* $W \in D^{(q,p)}([0,1], \mathbb{R}^{1+d})$ *has finitely many jumps. Let* $b \in C^{\beta,\gamma}$ *and let* $X, \widetilde{X} \in D([0,1], \mathbb{R}^m)$ *be given by*

$$\mathrm{d}X = b(X)^- \mathrm{d}W, \qquad \mathrm{d}\widetilde{X} = b(\widetilde{X}) \diamond \mathrm{d}W, \qquad X(0) = \widetilde{X}(0) = \xi.$$

Then

$$\|X - \widetilde{X}\|_{p\text{-var}} \le \|b\|_{\mathrm{Lip}} \|b\|_\infty K \sum_t |W(t) - W(t-)|^2,$$

where K depends only on $\|b\|_{C^{\beta,\gamma}}$, $\|W\|_{(q,p)\text{-var}}$, γ, β, p, and q, and the sum is over all jump times t of W.

Proof Let $t_1 < \cdots < t_n$ be the jump times of W; let $t_0 = 0$.

For $j \le n$, define X_j as the solution of forward RDE $\mathrm{d}X_j = b(X_j)^- \mathrm{d}W$, $X_j(0) = \xi$, on $[0, t_j]$, and as the solution of the geometric RDE $\mathrm{d}X_j = b(X_j) \diamond \mathrm{d}W$ on $[t_j, 1]$ with the initial condition taken from the solution on $[0, t_j]$.

For each j, the processes X_{j-1} and X_j coincide on $[0, t_j)$ but possibly differ at t_j. By Lemma 3.5 and the identity (3.6),

$$|X_j(t_j) - X_{j-1}(t_j)| \le \frac{1}{2}\|b\|_{\mathrm{Lip}}\|b\|_\infty |W(t_j) - W(t_j-)|^2. \tag{3.7}$$

On $[t_j, 1]$, both $X_{n,j-1}$ and $X_{n,j}$ solve the geometric RDE $\mathrm{d}X = b(X) \diamond \mathrm{d}W$, although with possibly different initial conditions. Recall that solutions of geometric RDEs are obtained from RDEs driven by continuous paths by inserting fictitious time intervals and linearly bridging the jumps. As such, they enjoy Lipschitz dependence on the initial condition (see [14, Theorem 12.10])

$$\begin{aligned} |\!|\!| X_j - X_{j-1} |\!|\!|_{p\text{-var};[t_j,1]} &= |X_j(t_j) - X_{j-1}(t_j)| + \|X_j - X_{j-1}\|_{p\text{-var};[t_j,1]} \\ &\le K|X_j(t_j) - X_{j-1}(t_j)|, \end{aligned} \tag{3.8}$$

where K depends only on $\|b\|_{C^{\beta,\gamma}}$, $\|W\|_{(q,p)\text{-var}}$, γ, β, p, and q.

It follows from (3.7) and (3.8) that

$$|\!|\!| X_j - X_{j-1} |\!|\!|_{p\text{-var}} \le \frac{1}{2}\|b\|_{\mathrm{Lip}}\|b\|_\infty K |W(t_j) - W(t_j-)|^2.$$

Observing that $X_0 = \widetilde{X}$ and $X_n = X$, and taking the sum over j, we obtain the result. □

Proof of Theorem 3.4 Denote by $\boldsymbol{\alpha}_{(q,p)\text{-var}}$ the metric on $D^{(q,p)\text{-var}}([0,1], \mathbb{R}^k)$ induced by the corresponding metric on $\mathscr{D}^{(q,p)\text{-var}}([0,1], \mathbb{R}^k)$ upon pairing paths with the linear path function ℓ_k, i.e. $\boldsymbol{\alpha}_{(q,p)\text{-var}}(X_1, X_2) = \boldsymbol{\alpha}_{(q,p)\text{-var}}((X_1, \ell_k), (X_2, \ell_k))$. Let $D^{0,(q,p)\text{-var}} \subset D^{(q,p)\text{-var}}$ denote the closure of smooth paths in $(D^{(q,p)\text{-var}}, \boldsymbol{\alpha}_{(q,p)\text{-var}})$. By the same argument as [7, Proposition 3.10 (v)], note that $D^{(q,p)\text{-var}} \subset D^{0,(q',p')\text{-var}}$ for all $q' > q$ and $p' > p$.

Fix $1 \le q' \le p' < 2$ with $p' \in (p, \gamma)$, $q' \in (q, \beta)$, and such that (3.1) holds with q, p replaced by q', p'. By Chevyrev and Friz [7, Proposition 2.9], convergence in $\mathcal{SM}_1$ is equivalent to convergence in $(D, \boldsymbol{\alpha}_\infty)$. By the Skorokhod representation theorem, we can thus suppose that a.s. $\lim_{n\to\infty} \boldsymbol{\alpha}_\infty(W_n, W) = 0$. Tightness of $\{\|W_n\|_{(q,p)\text{-var}}\}$ implies that a.s. there is a subsequence n_k such that $\limsup_{k\to\infty} \|W_{n_k}\|_{(q,p)\text{-var}} < \infty$, and thus $\|W\|_{(q,p)\text{-var}} < \infty$ a.s. by lower semi-continuity of (q, p)-variation. In addition, by a standard interpolation argument (cf. [7, Lemma 3.11]), it holds that $\boldsymbol{\alpha}_{(q',p')\text{-var}}(W_n, W) \to 0$ in probability, and therefore $W_n \to_w W$ in $(D^{0,(q',p')\text{-var}}, \boldsymbol{\alpha}_{(q',p')\text{-var}})$.

Since $(D^{0,(q',p')\text{-var}}, \boldsymbol{\alpha}_{(q',p')\text{-var}})$ is separable, we can again apply the Skorokhod representation theorem and suppose henceforth that, a.s., $W_n \to W$ in $\boldsymbol{\alpha}_{(q',p')\text{-var}}$ and $\sum |W_n(t) - W_n(t-)|^2 \to 0$ (we used here that $\sum |W_n(t) - W_n(t-)|^2$ converges in law to a constant).

An application of the continuity of solution map for generalised geometric RDEs (the proof of [7, Theorem 3.13] combined with [14, Theorem 12.10]; see Remark 3.2) shows that

$$((W_n, \widetilde{X}_n), \phi_b) \to ((W, X), \phi_b) \text{ in } (\mathscr{D}^{(q',p')\text{-var}}([0,1], \mathbb{R}^{1+d+m}), \boldsymbol{\alpha}_{(q',p')\text{-var}}), \tag{3.9}$$

where $\widetilde{X}_n$ solves the geometric RDE

$$\mathrm{d}\widetilde{X}_n = b(\widetilde{X}_n) \diamond \mathrm{d}W_n, \qquad \widetilde{X}_n(0) = \xi_n.$$

Furthermore, since clearly

$$\lim_{n\to\infty} \boldsymbol{\alpha}_\infty(((W_n, \widetilde{X}_n), \phi_b), ((W_n, \widetilde{X}_n), \ell_{1+d+m})) = 0, \tag{3.10}$$

it follows from [7, Lemma 3.11] that

$$\lim_{n\to\infty} \boldsymbol{\alpha}_{(q',p')\text{-var}}(((W_n, \widetilde{X}_n), \phi_b), ((W_n, \widetilde{X}_n), \ell_{1+d+m})) = 0. \tag{3.11}$$

It follows from Lemma 3.6 that $\lim_{n\to\infty} \|(W_n, \widetilde{X}_n) - (W_n, X_n)\|_{p'\text{-var}} = 0$, and in particular that $\boldsymbol{\sigma}_\infty((W_n, \widetilde{X}_n), (W_n, X_n)) \to 0$. By virtue of interpolation, for each $q'' > q'$ and $p'' > p'$, the identity map

$$(W, X) \mapsto ((W, X), \ell_{1+d+m}), \qquad \left(D^{(q',p')\text{-var}}, \boldsymbol{\sigma}_\infty\right) \to \left(D^{(q',p')\text{-var}}, \boldsymbol{\alpha}_{(q'',p'')\text{-var}}\right)$$

is uniformly continuous on sets bounded in (q', p')-variation (cf. [7, Proposition 3.12]), from which it follows that

$$\lim_{n\to\infty} \boldsymbol{\alpha}_{(q'',p'')\text{-var}}(((W_n, \widetilde{X}_n), \ell_{1+d+m}), ((W_n, X_n), \ell_{1+d+m})) = 0. \tag{3.12}$$

Combining (3.9), (3.11), and (3.12), we obtain

$$\lim_{n\to\infty} \boldsymbol{\alpha}_{(q'',p'')\text{-var}}(((W_n, X_n), \ell_{1+d+m}, ((W, X), \phi_b)) = 0.$$

Since $q'' > q' > q$ and $p'' > p' > p$ are arbitrary, the conclusion follows. □

We are now ready for the proof of Theorem 2.6.

Proof of Theorem 2.6 Defining the process $V_n : [0, \infty) \to [0, \infty)$, $V_n(t) = n^{-1}\lfloor tn \rfloor$, observe that X_n solves the forward RDE

$$\mathrm{d}X_n = a(X_n)^- \mathrm{d}V_n + b(X_n)^- \mathrm{d}W_n.$$

It follows from our assumptions that

$$(V_n, W_n) \to (\mathrm{id}, L) \quad \text{in the } \mathcal{SM}_1 \text{ topology} \tag{3.13}$$

and

$$\{\|(V_n, W_n)\|_{(1,p)\text{-var}}\}_{n\geq 1} \quad \text{is tight for every } p > \alpha'. \tag{3.14}$$

Furthermore, since $\alpha < 2$ and W_n makes at most n jumps of size at most $n^{-1/\alpha}\|v\|_\infty$,

$$\sum_t |W_n(t) - W_n(t-)|^2 \leq \|v\|_\infty^2 n^{1-2/\alpha} \to 0 \quad \text{as } n \to \infty. \tag{3.15}$$

Choose $p \in (\alpha', \gamma)$ and $q \in (1, \min\{p, \beta\})$ such that (3.1) is satisfied. By Theorem 3.4, it follows from (3.13), (3.14), and (3.15) that $\|L\|_{p\text{-var}} < \infty$ a.s. and

$$((V_n, W_n, X_n), \ell_{1+d+m}) \to_w ((\mathrm{id}, L, X), \phi_{(a,b)}) \tag{3.16}$$

in $(\mathscr{D}^{(q,p)\text{-var}}([0,1], \mathbb{R}^{1+d+m}), \boldsymbol{\alpha}_{(q,p)\text{-var}})$. Moreover, $\lim_{n\to 0}\|V_n - \mathrm{id}\|_{q\text{-var}} = 0$ and thus (3.16) readily implies that $((W_n, X_n), \ell_{d+m}) \to_w ((L, X), \phi_b)$ in $(\mathscr{D}^{p\text{-var}}([0,1], \mathbb{R}^{d+m}), \boldsymbol{\alpha}_{p\text{-var}})$. □

4 Results for Gibbs–Markov maps

In this section, we prove results on weak convergence to a Lévy process, and tightness in p-variation, for a class of uniformly expanding maps known as Gibbs–Markov maps [1]. The weak convergence result extends work of [1,22,33,42] from scalar-valued observables to $\mathbb{R}^d$-valued observables. The result on tightness in p-variation is new even for $d = 1$.

4.1 Gibbs–Markov maps

Let (Z, d) be a bounded metric space with Borel sigma-algebra $\mathcal{B}$ and finite Borel measure ν, and an at most countable partition $\mathcal{P}$ of Z (up to a zero measure set) with $\nu(a) > 0$ for each $a \in \mathcal{P}$. Let $F\colon Z \to Z$ be a nonsingular ergodic measurable transformation. We assume that F is a *Gibbs–Markov map*. That is, there are constants $\lambda > 1$, $K > 0$ and $\theta \in (0, 1]$ such that for all $z, z' \in a$ and $a \in \mathcal{P}$:

- Fa is a union of partition elements and F restricts to a (measure-theoretic) bijection from a to Fa; moreover $\inf_{a\in\mathcal{P}} \nu(Fa) > 0$;
- $d(Fz, Fz') \geq \lambda d(z, z')$;
- the inverse Jacobian $\zeta_a = \frac{d\nu}{d\nu\circ F}$ of the restriction $F\colon a \to Fa$ satisfies

$$\left|\log \zeta_a(z) - \log \zeta_a(z')\right| \leq K d(Fz, Fz')^\theta. \tag{4.1}$$

It is standard (see for example [1, Corollary p. 199]) that there is a unique F-invariant probability measure μ_Z absolutely continuous with respect to ν, with bounded density

$d\mu_Z/d\nu$. The measure μ_Z is ergodic and we suppose for simplicity that μ_Z is mixing. (The nonmixing case is also covered by standard arguments, see for example the end of the proof of [33, Proposition 4.3], but is not required here.)

Definition 4.1 We say that an $\mathbb{R}^d$-valued random variable ξ is *regularly varying* with index $\alpha > 0$ if there exists a probability measure σ on $\mathcal{B}(\mathbb{S}^{d-1})$, the Borel sigma-algebra on the unit sphere $\mathbb{S}^{d-1} = \{x \in \mathbb{R}^d : |x| = 1\}$, such that

$$\lim_{t\to\infty} \frac{\mathbb{P}(|\xi| > rt,\ \xi/|\xi| \in B)}{\mathbb{P}(|\xi| > t)} = r^{-\alpha}\sigma(B)$$

for all $r > 0$ and $B \in \mathcal{B}(\mathbb{S}^{d-1})$ with $\sigma(\partial B) = 0$.

Recall that an α-stable random variable X in $\mathbb{R}^d$ with $\alpha \in (1, 2)$ and $\mathbb{E}\, X = 0$ has characteristic function

$$\mathbb{E}\exp(iu\cdot X) = \exp\left\{-\int_{\mathbb{S}^{d-1}} |u\cdot s|^\alpha\Big(1 - i\,\mathrm{sgn}(u\cdot s)\tan\frac{\pi\alpha}{2}\Big)\mathrm{d}\Lambda(s)\right\}, \quad u \in \mathbb{R}^d.$$

Here Λ is a finite nonnegative Borel measure on $\mathbb{S}^{d-1}$ with $\Lambda(\mathbb{S}^{d-1}) > 0$, known as the *spectral measure* [39, Section 2.3]. It is a direct verification that γX, with $\gamma \ge 0$, has spectral measure $\gamma^\alpha\Lambda$.

We say that an α-stable Lévy process L_α has spectral measure Λ if $L_\alpha(1)$ has spectral measure Λ.

Fix a function $\tau\colon Z \to \{1, 2, \ldots\}$ that is constant on each $a \in \mathcal{P}$ with value $\tau(a)$ such that $\int_Z \tau\,\mathrm{d}\mu_Z < \infty$. Let $V\colon Z \to \mathbb{R}^d$ be integrable with $\int_Z V\,\mathrm{d}\mu_Z = 0$. Assume that there exists $C_0 > 0$ such that for and all $z, z' \in a$, $a \in \mathcal{P}$,

$$|V(z)| \le C_0\tau(a) \qquad \text{and} \qquad |V(z) - V(z')| \le C_0\tau(a)d(Fz, Fz')^\theta. \tag{4.2}$$

Suppose that b_n is a sequence of positive numbers and define the càdlàg process

$$W_n(t) = b_n^{-1}\sum_{j=0}^{\lfloor nt\rfloor - 1} V\circ F^j.$$

We consider W_n as a random element on the probability space (Z, μ_Z). Throughout this section, $\|\cdot\|_p$ denotes the L^p norm on (Z, μ_Z) for $1 \le p \le \infty$ and $\mathbb{E}$ denotes expectation with respect to μ_Z.

We now state the main results of this section.

Theorem 4.2 *Suppose that*

- *V is regularly varying on (Z, μ_Z) with index $\alpha \in (1, 2)$ and σ as in Definition 4.1,*
- *b_n satisfies $\lim_{n\to\infty} n\mu_Z(|V| > b_n) = 1$,*
- *$V - \mathbb{E}(V \mid \mathcal{P}) \in L^p$ for some $p > \alpha$.*

Then $W_n \to_w L_\alpha$ in the $\mathcal{SJ}_1$ topology as $n \to \infty$, where L_α is the α-stable Lévy process with spectral measure $\Lambda = \cos\frac{\pi\alpha}{2}\Gamma(1-\alpha)\sigma$.

Remark 4.3

(a) If V is regularly varying and $\lim_{n\to\infty} n\mu_Z(|V| > b_n) = 1$, then b_n is a regularly varying sequence. In particular, if $\mu_Z(|V| > n) \sim cn^{-\alpha}$ for some $c > 0$, then $b_n \sim c^{1/\alpha} n^{1/\alpha}$.

(b) In many examples (including the intermittent maps in Sect. 6.2), $\tau \in L^q$ for each $q < \alpha$, and there exist $C > 0$ and $\beta \in (0, 1)$ such that $|V(z) - V(z')| \le C\tau^\beta$ for all $z, z' \in a$, $a \in \mathcal{P}$. This implies that $V - \mathbb{E}(V \mid \mathcal{P}) \in L^p$ for some $p > \alpha$.

Theorem 4.4 *Suppose that τ is regularly varying with index $\alpha \in (1, 2)$ on (Z, μ_Z), and that b_n satisfies $\lim_{n\to\infty} n\mu_Z(\tau > b_n) = 1$. Then $\sup_n \int_Z \|W_n\|_{p\text{-var}} d\mu_Z < \infty$ for all $p > \alpha$.*

4.2 Preliminaries about Gibbs–Markov maps

We recall the following standard result.

Lemma 4.5 *Let $V : Z \to \mathbb{R}^d$ be integrable with $\int_Z V d\mu_Z = 0$ and satisfying (4.2). Then*

(a) *$V = m + \chi \circ F - \chi$, where m is integrable with $\mathbb{E}(m \mid F^{-1}\mathcal{B}) = 0$, and $\|\chi\|_\infty \le CC_0$ with $C > 0$ independent of V.*

(b) *For every $p \in (1, 2]$ there is a constant $C(p)$, depending only on p, such that*

$$\Big\|\max_{k\le n}\Big|\sum_{j=0}^{k-1} V \circ F^j\Big|\Big\|_p \le C(p)n^{1/p}(\|\chi\|_\infty + \|V\|_p).$$

(We do not exclude the case $\|V\|_p = \infty$.)

Proof For $z, z' \in Z$, let $s(z, z')$ be the *separation time*, i.e. the minimal nonnegative integer such that $F^{s(z,z')}(z)$ and $F^{s(z,z')}(z')$ belong to different elements of $\mathcal{P}$. Let d_θ be the separation metric on Z:

$$d_\theta(z, z') = \lambda^{-\theta s(z,z')}.$$

Note that $d(z, z')^\theta \le d_\theta(z, z')(\operatorname{diam} Z)^\theta$, so θ-Hölder observables with respect to d are d_θ-Lipschitz. For an observable $\phi : Z \to \mathbb{R}^d$, let

$$\|\phi\| = \|\phi\|_\infty + \sup_{z\ne z'} \frac{|\phi(z) - \phi(z')|}{d_\theta(z, z')}.$$

Let $P : L^1(\mu_Z) \to L^1(\mu_Z)$ be the transfer operator corresponding to F and μ_Z, i.e. $\int_Z P\phi\, w d\mu_Z = \int_Z \phi\, w \circ F d\mu_Z$ for all $\phi \in L^1$, $w \in L^\infty$.

By for example [1, Section 1], there are constants $C_1 > 0$, $\gamma \in (0, 1)$ such that $\|P^k\phi\| \le C_1\gamma^k\|\phi\|$ for all $\phi : Z \to \mathbb{R}^d$ with $\mathbb{E}\,\phi = 0$ and all $k \ge 0$.

By Melbourne and Nicol [30, Lemma 2.2], there is a constant $C_2 > 0$ independent of V such that $\|PV\| \le C_0C_2$ for all V satisfying the stated conditions. Hence

$$\|P^kV\| = \|P^{k-1}PV\| \le C_1\gamma^{k-1}\|PV\| \le C_0C_1C_2\gamma^{k-1}.$$

Let $\chi = \sum_{k=1}^{\infty} P^kV$. Then $\|\chi\|_\infty \le \|\chi\| \le C_0C_1C_2(1-\gamma)^{-1}$. Let $m = V - \chi \circ F + \chi$. Define $U\colon L^1(\mu_Z) \to L^1(\mu_Z)$ by $U\phi = \phi \circ F$. Then $PU = I$ and $UP = \mathbb{E}(\cdot \mid F^{-1}\mathcal{B})$. Hence $\mathbb{E}(m \mid F^{-1}\mathcal{B}) = UPm = U(PV - \chi + P\chi) = 0$ proving part (a).

For part (b), we proceed as in the proof of [33, Proposition 4.3]. Fix $n > 0$ and let $M_k^n = \sum_{j=n-k}^{n-1} m \circ F^j$. By (a), M_k^n is a martingale on $0 \le k \le n$. By Burkholder's inequality, there is a constant $C(p)$ depending only on p such that

$$\Big\|\max_{k\le n}|M_k^n|\Big\|_p \le C(p)n^{1/p}\|m\|_p \le C(p)n^{1/p}(2\|\chi\|_\infty + \|V\|_p).$$

Next,

$$\bigg\|\max_{k\le n}\bigg|\sum_{j=0}^{k-1} V \circ F^j\bigg|\bigg\|_p \le 2\|\chi\|_\infty + 2\Big\|\max_{k\le n}|M_k^n|\Big\|_p,$$

and part (b) follows. □

For sigma-algebras $\mathcal{F}$ and $\mathcal{G}$ on a common probability space $(\Omega, \mathbb{P})$, define

$$\psi(\mathcal{F}, \mathcal{G}) = \sup\left\{\frac{\big|\mathbb{P}(A \cap B) - \mathbb{P}(A)\,\mathbb{P}(B)\big|}{\mathbb{P}(A)\,\mathbb{P}(B)} : A \in \mathcal{F},\ \ B \in \mathcal{G}\right\}.$$

For $0 \le n \le k$, let $\mathcal{P}_n^k$ be the smallest sigma-algebra which contains $F^{-j}\mathcal{P}$ for $j = n, \ldots, k$. A standard property of mixing Gibbs–Markov maps (see for example [1, Section 1]) is that there exist $\gamma \in (0, 1)$ and $C > 0$ such that for all $k \ge 0$, $n \ge 1$,

$$\psi\big(\mathcal{P}_0^k, \mathcal{P}_{n+k}^\infty\big) \le C\gamma^n, \tag{4.3}$$

where the probability measure in the definition of ψ is μ_Z.

4.3 Weak convergence to a Lévy process

In this subsection, we prove Theorem 4.2. We use the following result due to Tyran-Kamińska [41].

Theorem 4.6 *Let $X_0, X_1, \ldots$ be a strictly stationary sequence of integrable $\mathbb{R}^d$-valued random variables with $\mathbb{E}\,X_0 = 0$. For $0 \le n \le k$, let $\mathcal{F}_n^k$ denote the sigma-algebra generated by $\{X_n, \ldots, X_k\}$. Suppose that:*

(a) *X_0 is regularly varying with index $\alpha \in [1, 2)$ and σ as in Definition 4.1.*

(b) $\sum_{j\ge 0} \psi(2^j) < \infty$, *where* $\psi(n) = \sup_{k\ge 0} \psi(\mathcal{F}_0^k, \mathcal{F}_{n+k}^\infty)$.
(c) $\lim_{n\to\infty} \mathbb{P}\big(|X_j| > \epsilon b_n \bigm| |X_0| > \epsilon b_n\big) = 0$ *for all* $\epsilon > 0$ *and* $j \ge 1$, *where the sequence* b_n *is such that* $\lim_{n\to\infty} n\,\mathbb{P}(|X_0| > b_n) = 1$.

Then as $n \to \infty$, *the random process* W_n *given by* $W_n(t) = b_n^{-1} \sum_{j=0}^{\lfloor nt\rfloor -1} X_j$ *converges to an* α*-stable Lévy process* L_α *in* $D([0,1], \mathbb{R}^d)$ *in the* $\mathcal{SJ}_1$ *topology.*

Remark 4.7 It is implicit in [41] that L_α has spectral measure $\Lambda = \cos\frac{\pi\alpha}{2}\Gamma(1-\alpha)\sigma$, where σ is the measure on $\mathbb{S}^{d-1}$ for X_0 as in Definition 4.1.

Proof of Theorem 4.6 We verify the hypotheses of [41, Theorem 1.1]. In the notation of [41], observe that (b) and [41, Lemma 4.8] together with $\rho \le \psi$ imply that [41, Eq. (1.6)] holds. Moreover, (c) and [41, Corollary 1.3] together with $\varphi \le \psi$ imply that [41, $\mathbf{LD}(\phi_0)$] holds (for inequalities concerning ρ, ψ, and φ, see [4]). $\square$

Write $V = V' + V''$ where $V' = \mathbb{E}(V \mid \mathcal{P})$. Let

$$W_n'(t) = b_n^{-1} \sum_{j=0}^{\lfloor nt\rfloor -1} V' \circ F^j, \qquad W_n''(t) = b_n^{-1} \sum_{j=0}^{\lfloor nt\rfloor -1} V'' \circ F^j.$$

Proposition 4.8 (i) W_n' *converges in* $\mathcal{SJ}_1$ *to the* α*-stable Lévy process* L_α *with spectral measure* $\Lambda = \cos\frac{\pi\alpha}{2}\Gamma(1-\alpha)\sigma$.
(ii) $\big\|\sup_{t\in[0,1]} |W_n''(t)|\big\|_1 \to 0$ *as* $n \to \infty$.

Proof To prove part (i), we verify the hypotheses of Theorem 4.6 with $X_k = V' \circ F^k$. Since μ_Z is F-invariant, $\{V' \circ F^k\}_{k\ge 0}$ is a strictly stationary sequence of $\mathbb{R}^d$-valued random variables. The remaining hypotheses are verified as follows

(a) The observable V is regularly varying with index α and measure σ, and $V'' \in L^p$ with $p > \alpha$, so $V' = V - V''$ is regularly varying with the same α and σ.
(b) This is a consequence of (4.3).
(c) It follows from (4.3) and invariance of μ_Z under F that

$$\mu_Z\big(|V' \circ F^j| > \epsilon b_n \bigm| |V'| > \epsilon b_n\big) \lesssim \mu_Z(|V'| > \epsilon b_n).$$

Now we prove part (ii). By the assumptions of Theorem 4.2, $V'' \in L^p$ for some $p \in (\alpha, 2)$. Note that $|V''| \lesssim \tau$, $\mathbb{E}\,V'' = 0$ and for each $z, z' \in a$, $a \in \mathcal{P}$,

$$|V''(z) - V''(z')| = |V(z) - V(z')| \le C_0 \tau(a) d(Fz, Fz')^\theta .$$

Hence by Lemma 4.5(b), $\big\|\max_{k\le n} |\sum_{j=0}^{k-1} V'' \circ F^j|\big\|_p \lesssim n^{1/p} = o(b_n)$.

Proof of Theorem 4.2 By Proposition 4.8, $W_n = W_n' + W_n'' \to_w L_\alpha$. $\square$

4.4 Tightness in *p*-variation

In this subsection we prove Theorem 4.4.

First we record the following elementary properties of τ. (The Gibbs–Markov structure is not required here; the proof only uses that τ is regularly varying with values in $\{1, 2, \ldots\}$ and that μ_Z is F-invariant.)

Proposition 4.9 *Let $p > \alpha$. Then*

(a) $\mathbb{E}(\tau^p 1_{\{\tau \le b_n\}}) = O(n^{-1} b_n^p)$,
(b) $\mathbb{E}(\tau 1_{\{\tau \ge b_n\}}) = O(n^{-1} b_n)$,
(c) $\mathbb{E}\big\{\big(\sum_{j=0}^{n-1} \tau^p \circ F^j\big)^{1/p}\big\} = O(b_n)$.

Proof We have

$$\mathbb{E}\left(\tau^p 1_{\{\tau \le b_n\}}\right) = \sum_{j \le b_n} j^p \mu_Z(\tau = j) \le \sum_{j \le b_n} (j^p - (j-1)^p)\mu_Z(\tau \ge j)$$
$$\le p \sum_{j \le b_n} j^{p-1} \mu_Z(\tau \ge j).$$

By Karamata's theorem [3, Proposition 1.5.8], $\mathbb{E}(\tau^p 1_{\{\tau \le b_n\}}) \lesssim b_n^p \mu_Z(\tau \ge b_n)$, so part (a) follows by definition of b_n. A similar calculation proves part (b). Next,

$$\left(\sum_{j=0}^{n-1} \tau^p \circ F^j\right)^{1/p} \le \left(\sum_{j=0}^{n-1} \left(\tau^p 1_{\{\tau > b_n\}}\right) \circ F^j\right)^{1/p} + \left(\sum_{j=0}^{n-1} \left(\tau^p 1_{\{\tau \le b_n\}}\right) \circ F^j\right)^{1/p}$$
$$\le \sum_{j=0}^{n-1} \left(\tau 1_{\{\tau > b_n\}}\right) \circ F^j + \left(\sum_{j=0}^{n-1} \left(\tau^p 1_{\{\tau \le b_n\}}\right) \circ F^j\right)^{1/p}.$$

By Jensen's inequality, invariance of μ_Z and parts (a) and (b),

$$\mathbb{E}\left\{\left(\sum_{j=0}^{n-1} \tau^p \circ F^j\right)^{1/p}\right\}$$
$$\le \sum_{j=0}^{n-1} \mathbb{E}\big\{\left(\tau 1_{\{\tau > b_n\}}\right) \circ F^j\big\} + \left(\sum_{j=0}^{n-1} \mathbb{E}\big\{\left(\tau^p 1_{\{\tau \le b_n\}}\right) \circ F^j\big\}\right)^{1/p}$$
$$= n\, \mathbb{E}(\tau 1_{\{\tau > b_n\}}) + \big(n\, \mathbb{E}(\tau^p 1_{\{\tau \le b_n\}})\big)^{1/p} \lesssim b_n,$$

proving part (c). □

Write $V = V_n' - \mathbb{E}\, V_n' + V_n''$, where

$$V_n' = V 1_{\{\tau > b_n\}}, \quad V_n'' = V 1_{\{\tau \le b_n\}} - \mathbb{E}(V 1_{\{\tau \le b_n\}}).$$

Accordingly, define $W_n = W_n' - \mathbb{E}\, W_n' + W_n''$, where

$$W_n'(t) = b_n^{-1} \sum_{j=0}^{\lfloor nt \rfloor -1} V_n' \circ F^j, \qquad W_n''(t) = b_n^{-1} \sum_{j=0}^{\lfloor nt \rfloor -1} V_n'' \circ F^j.$$

Proposition 4.10 $\sup_n \mathbb{E}\, \|W_n'\|_{1\text{-var}} < \infty$.

Proof By Proposition 4.9(b), $\mathbb{E}\,|V_n'| \le C_0\, \mathbb{E}\big(\tau 1_{\{\tau > b_n\}}\big) \lesssim n^{-1} b_n$. Hence

$$\mathbb{E}\, \|W_n'\|_{1\text{-var}} = \mathbb{E} \left(b_n^{-1} \sum_{j=0}^{n-1} |V_n'| \circ F^j \right) = n b_n^{-1}\, \mathbb{E}\,|V_n'| = O(1),$$

as required. □

Proposition 4.11 $\sup_n \mathbb{E}\, \|W_n''\|_{p\text{-var}}^p < \infty$ *for all* $p \in (\alpha, 2)$.

Proof Note that $\mathbb{E}\, V_n'' = 0$, that $|V_n''| \le |V| + \mathbb{E}\,|V| \le C_1 \tau$ where $C_1 = C_0 + \mathbb{E}\,|V|$, and that $|V_n''(z) - V_n''(z')| \le |V(z) - V(z')| \le C_0 \tau(a) d(Fz, Fz')^\theta$ for all $z, z' \in a$, $a \in \mathcal{P}$. By Lemma 4.5(a), $V_n'' = m_n + \chi_n \circ F - \chi_n$, where $\sup_n \|\chi_n\|_\infty < \infty$ and $\mathbb{E}(m_n \mid F^{-1}\mathcal{B}) = 0$. Then

$$\|m_n\|_p \le \|V_n''\|_p + 2\|\chi_n\|_p \le 2\|V 1_{\{\tau \le b_n\}}\|_p + 2\|\chi_n\|_\infty$$

and $\mathbb{E}\,|V 1_{\{\tau \le b_n\}}|^p \le C_0^p\, \mathbb{E}\left(\tau^p 1_{\{\tau \le b_n\}}\right) \lesssim n^{-1} b_n^p$ by Proposition 4.9(a). The assumptions of Theorem 4.4 imply that $b_n^p \gtrsim n$. Hence

$$\mathbb{E}\,|m_n|^p \lesssim n^{-1} b_n^p. \tag{4.4}$$

Write $W_n'' = M_n + B_n$ where

$$\begin{aligned} M_n(t) = b_n^{-1} \sum_{j=0}^{\lfloor nt \rfloor -1} m_n \circ F^j, \quad B_n(t) &= b_n^{-1} \sum_{j=0}^{\lfloor nt \rfloor -1} (\chi_n \circ F - \chi_n) \circ F^j \\ &= b_n^{-1}(\chi_n \circ F^{\lfloor nt \rfloor} - \chi_n). \end{aligned}$$

Let $M_n^-(t) = b_n^{-1} \sum_{j=1}^{\lfloor nt \rfloor} m_n \circ F^{n-j}$. Then M_n^- is a martingale since $\mathbb{E}(m_n \mid F^{-1}\mathcal{B}) = 0$. By [36, Theorem 2.1] and (4.4),

$$\mathbb{E}\, \|M_n\|_{p\text{-var}}^p = \mathbb{E}\, \|M_n^-\|_{p\text{-var}}^p \lesssim b_n^{-p} \sum_{j=1}^{n} \mathbb{E}\,|m_n \circ F^{n-j}|^p = n b_n^{-p}\, \mathbb{E}\,|m_n|^p \lesssim 1. \tag{4.5}$$

Finally, $\|B_n\|_{p\text{-var}}^p \le b_n^{-p}\, n\, (2\|\chi_n\|_\infty)^p \lesssim n b_n^{-p} \lesssim 1$ for $p > \alpha$.

Remark 4.12 For our purposes, it is sufficient to control the first moment $\mathbb{E}\,\|W_n''\|_{p\text{-var}}$. Hence we could have used the simpler result [26, Proposition 2] in place of the sharp result [36, Theorem 2.1]; this would give $\sup_n \mathbb{E}\,\|W_n''\|_{p\text{-var}}^q < \infty$ for all $p > \alpha$ and $q < p$.

Proof of Theorem 4.4 Combine Propositions 4.10 and 4.11. □

5 Inducing weak convergence and tightness in *p*-variation

A general principle in smooth ergodic theory is that limit laws for dynamical systems are often inherited from the corresponding laws for a suitable induced system [18,20,31,33,38]. In this section, we show that this principle applies to weak convergence in $D([0,1],\mathbb{R}^d)$ with the $\mathcal{SM}_1$ topology and to tightness in p-variation. The results hold in a purely probabilistic setting.

Let Y be a measurable space and $f\colon Y \to Y$ a measurable transformation. Suppose that $Z \subset Y$ is a measurable subset with a measurable return time $\tau\colon Z \to \{1,2,\ldots\}$, i.e. $f^{\tau(z)}(z) \in Z$ for each $z \in Z$. (It is not assumed that τ is the first return time.) Define the induced map

$$F\colon Z \to Z, \qquad Fz = f^{\tau(z)}(z).$$

Suppose that μ_Z is an ergodic F-invariant probability measure and that $\bar\tau = \int_Z \tau \, d\mu_Z < \infty$.

Define the tower $f_\Delta\colon \Delta \to \Delta$

$$\Delta = \{(z,\ell) : z \in Z, 0 \le \ell < \tau(z)\}, \qquad f_\Delta(z,\ell) = \begin{cases} (z,\ell+1), & \ell < \tau(z)-1, \\ (Fz,0), & \ell = \tau(z)-1, \end{cases} \tag{5.1}$$

with ergodic f_Δ-invariant probability measure $\mu_\Delta = (\mu_Z \times \text{counting})/\bar\tau$. The map $\pi\colon \Delta \to Y$, $\pi(z,\ell) = f^\ell z$ defines a measurable semiconjugacy between f_Δ and f, so $\mu = \pi_*\mu_\Delta$ is an ergodic f-invariant probability measure on Y.

It is convenient to identify Z with $Z \times \{0\} \subset \Delta$. Then on the tower, τ is the first return time to Z.

Let $v\colon Y \to \mathbb{R}^d$ be measurable and define the corresponding *induced observable*

$$V\colon Z \to \mathbb{R}^d, \qquad V(z) = \sum_{j=0}^{\tau(z)-1} v(f^j z). \tag{5.2}$$

Let $v_k = \sum_{j=0}^{k-1} v \circ f^j$. To measure how well the excursion $\{v_k(z)\}_{0\le k\le \tau(z)}$ approximates the straight and monotone path from 0 to $V(z)$, we define $V^*\colon Z \to \mathbb{R}^d$,

$$V^* = \inf_{c\in\mathbb{R}^d, |c|=1} \Big(\max_{0\le k\le \ell\le\tau} c\cdot\big(v_k - v_\ell\big) + \max_{0\le k\le\tau} \big|v_k - (c\cdot v_k)c\big| \Big). \tag{5.3}$$

Note that $V^*(z) = 0$ if and only if there exist $0 = s_0 \le s_1 \le \cdots \le s_{\tau(z)} = 1$ such that $v_k(z) = s_k V(z)$ for $0 \le k \le \tau(z)$.

Let b_n be a sequence of positive numbers, bounded away from 0, and define

$$W_n(t) = b_n^{-1} \sum_{j=0}^{\lfloor nt \rfloor - 1} v \circ f^j \quad \text{and} \quad \widetilde{W}_n(t) = b_n^{-1} \sum_{j=0}^{\lfloor nt \rfloor - 1} V \circ F^j. \tag{5.4}$$

In this section, the notation $\to_\mu$ and $\to_{\mu_Z}$ is used to denote weak convergence for random variables defined on the probability spaces (Y, μ) and (Z, μ_Z) respectively. We prove:

Theorem 5.1 *Suppose that $\widetilde{W}_n \to_{\mu_Z} \widetilde{W}$ in the $\mathcal{SM}_1$ topology for some random process $\widetilde{W}$. Suppose further that*

$$b_n^{-1} \max_{k<n} V^* \circ F^k \to_{\mu_Z} 0.$$

Then $W_n \to_\mu W$ in the $\mathcal{SM}_1$ topology where $W(t) = \widetilde{W}(t/\bar{\tau})$.

Theorem 5.2 *Suppose that τ is regularly varying with index $\alpha > 1$ on (Z, μ_Z), and that b_n satisfies $\lim_{n\to\infty} n\mu_Z(\tau > b_n) = 1$. Let $v \in L^\infty$. Suppose that the family of random variables $\|\widetilde{W}_n\|_{p\text{-var}}$ is tight on (Z, μ_Z) for some $p > \alpha$. Then the family $\|W_n\|_{p\text{-var}}$ is tight on (Y, μ).*

Remark 5.3 The assumptions of Theorem 5.2 on τ can be relaxed. If $\tau' \colon Z \to \{1, 2, \ldots\}$ is regularly varying with index $\alpha > 1$ on (Z, μ_Z) and b_n satisfies $\lim_{n\to\infty} n\mu_Z(\tau' > b_n) = 1$, then the result holds for all $\tau \le \tau'$.

5.1 Inducing convergence in $\mathcal{SM}_1$ topology

In this subsection, we prove Theorem 5.1. Our proof closely follows the analogous proof in [33], with the difference that we work in $\mathbb{R}^d$ instead of $\mathbb{R}$.

Since $\pi : \Delta \to Y$ is a measure-preserving semiconjugacy, we may suppose without loss of generality that $Y = \Delta$ and $f = f_\Delta$ as in (5.1). In particular, we may suppose that τ is the first return time.

Define

$$u \colon Y \to \mathbb{R}^d, \qquad u(y) = \begin{cases} V(z), & y = (z, \tau(z) - 1), \\ 0, & \text{otherwise.} \end{cases}$$

Let

$$U_n(t) = b_n^{-1} \sum_{j=0}^{\lfloor nt \rfloor - 1} u \circ f^j.$$

Thus defined, the restriction of U_n to Z corresponds to U_n in [33].

Lemma 5.4 *$U_n \to_{\mu_Z} W$ in the $\mathcal{SM}_1$ topology.*

Proof For the case $d = 1$, see [33, Lemma 3.4]. The proof for all $d \geq 1$ goes through unchanged. □

Next we control *excursions:* we estimate the distance between U_n and W_n in the $\mathcal{SM}_1$ topology.

Proposition 5.5 *Let $w \in D([T_0, T_1], \mathbb{R}^d)$ and define $\phi\colon [T_0, T_1] \to \mathbb{R}^d$ to be the linear path with $\phi(T_0) = w(T_0)$ and $\phi(T_1) = w(T_1)$. Then for each $c \in \mathbb{R}^d$ with $|c| = 1$,*

$$d_{\mathcal{SM}_1}(w, \phi) \leq T_1 - T_0 + 2 \sup_{T_0 \leq s < t \leq T_1} c \cdot w(t, s) + 2 \sup_{T_0 \leq t \leq T_1} \big| w(T_0, t) - (c \cdot w(T_0, t))c \big|,$$

where $w(a, b) = w(b) - w(a)$.

Proof Without loss of generality, we suppose that $w(T_0) = 0$. Define $\chi : [T_0, T_1] \to [0, \infty)$ and $\psi\colon [T_0, T_1] \to \mathbb{R}^d$ to be $\chi(t) = \sup_{s \leq t} c \cdot w(s)$ and $\psi(t) = \chi(t)c$. Then ψ is a monotone path in the direction of c.

Observe that $|w(t) - \psi(t)| \leq \chi(t) - c \cdot w(t) + |w(t) - (c \cdot w(t))c|$. Hence

$$\sup_t |w(t) - \psi(t)| \leq \sup_{s<t} c \cdot w(t, s) + \sup_t |w(t) - (c \cdot w(t))c|. \tag{5.5}$$

Further, let $\xi\colon [T_0, T_1] \to \mathbb{R}^d$ be the linear path with $\xi(T_0) = w(T_0) = 0$ and $\xi(T_1) = \psi(T_1) = \chi(T_1)c$. Since ξ is a reparametrisation of ψ (up to linear jumps),

$$d_{\mathcal{SM}_1}(\xi, \psi) \leq T_1 - T_0. \tag{5.6}$$

Also, for each $\epsilon > 0$ there is $s \in [T_0, T_1]$ such that $|\chi(T_1) - c \cdot w(s)| \leq \epsilon$. Then

$$\begin{aligned} \sup_t |\phi(t) - \xi(t)| = |\phi(T_1) - \xi(T_1)| &\leq |w(T_1) - (c \cdot w(s))c| + \epsilon \\ &\leq |w(T_1) - (c \cdot w(T_1))c| + c \cdot (w(s) - w(T_1)) + \epsilon. \end{aligned} \tag{5.7}$$

The result follows from (5.5), (5.6), (5.7) and that ϵ can be taken arbitrarily small. □

For $s \leq t$, let $d_{\mathcal{SM}_1,[s,t]}$ denote the distance on $[s, t]$. Let $\tau_k = \sum_{j=0}^{k-1} \tau \circ F$.

Corollary 5.6 *For each n and k, on Z,*

$$d_{\mathcal{SM}_1,[0,\tau_k/n]}(U_n, W_n) \leq 2 \max_{0\leq j<k} \left\{ \frac{\tau \circ F^j}{n} + \frac{V^* \circ F^j}{b_n} \right\}.$$

Proof Denote $T_j = \tau_j/n$. Since we restrict to Z, each interval $[T_j, T_{j+1}]$, including with $j = 0$, corresponds to a complete excursion with $U_n(T_j) = W_n(T_j)$ and $U_n(T_{j+1}) = W_n(T_{j+1})$. Fix j and let $\phi\colon [T_j, T_{j+1}] \to \mathbb{R}^d$ be the linear path such that

$\phi(T_j) = U_n(T_j)$ and $\phi(T_{j+1}) = U_n(T_{j+1})$. Recall that U_n is constant on $[T_j, T_{j+1})$. By Proposition 5.5,

$$d_{\mathcal{SM}_1,[T_j,T_{j+1}]}(U_n, \phi) \le T_{j+1} - T_j,$$
$$d_{\mathcal{SM}_1,[T_j,T_{j+1}]}(W_n, \phi) \le T_{j+1} - T_j + \frac{2}{b_n} V^* \circ F^j.$$

Hence

$$d_{\mathcal{SM}_1,[T_j,T_{j+1}]}(U_n, W_n) \le 2(T_{j+1} - T_j) + \frac{2}{b_n} V^* \circ F^j = \frac{2}{n} \tau \circ F^j + \frac{2}{b_n} V^* \circ F^j.$$

Finally,

$$d_{\mathcal{SM}_1,[0,T_k]}(U_n, W_n) \le \max_{j<k} d_{\mathcal{SM}_1,[T_j,T_{j+1}]}(U_n, W_n)\ ,$$

and the result follows. □

Lemma 5.7 $d_{\mathcal{SM}_1,[0,T]}(U_n, W_n) \to_{\mu_Z} 0$ *for all* $T > 0$.

Proof Fix $T > 0$ and define the random variables $k = k(n) = \max\{j \ge 0 : \tau_j/n \le T\}$ on Z. Consider the processes U_n, W_n on Z, where the time interval $[0, \tau_k/n]$ corresponds to k complete excursions, while $[\tau_k/n, T]$ is the final incomplete excursion.

By Corollary 5.6 and the assumptions of Theorem 5.1,

$$d_{\mathcal{SM}_1,[0,\tau_k/n]}(U_n, W_n) \le 2 \max_{j<k} \left\{ \frac{\tau \circ F^j}{n} + \frac{V^* \circ F^j}{b_n} \right\} \to_{\mu_Z} 0.$$

For $y = (z, \ell) \in Y$, let $E(y) = \sum_{j=0}^{\tau(z)-1} \left| v(f^j z) \right|$. Since μ is f-invariant and $b_n \to \infty$, we have $b_n^{-1} E \circ f^{\lfloor nT \rfloor} \to_\mu 0$. Since μ_Z is absolutely continuous with respect to μ, we also have $b_n^{-1} E \circ f^{\lfloor nT \rfloor} \to_{\mu_Z} 0$. Hence

$$\begin{aligned} d_{\mathcal{SM}_1,[0,T]}(U_n, W_n) &\le d_{\mathcal{SM}_1,[0,\tau_k/n]}(U_n, W_n) + \sup_{[\tau_k/n,T]} |U_n - W_n| \\ &\le d_{\mathcal{SM}_1,[0,\tau_k/n]}(U_n, W_n) + \frac{1}{b_n} E \circ f^{\lfloor nT \rfloor} \to_{\mu_Z} 0 \end{aligned}$$

as required. □

Proof of Theorem 5.1 By Lemma 5.7, $d_{\mathcal{SM}_1,[0,T]}(U_n, W_n) \to_{\mu_Z} 0$ for every T. By Lemma 5.4, $U_n \to_{\mu_Z} W$ in $\mathcal{SM}_1$. Hence $W_n \to_{\mu_Z} W$ in $\mathcal{SM}_1$. The required convergence of $W_n \to_\mu W$ in $\mathcal{SM}_1$ follows from strong distributional convergence [47, Theorem 1] upon verifying that $d_{\mathcal{SM}_1}(W_n, W_n \circ f) \le d_{\mathcal{SJ}_1}(W_n, W_n \circ f) \to_\mu 0$ in the same way as [47, Corollary 3]. □

5.2 Inducing tightness in p-variation

In this subsection we prove Theorem 5.2. Again, we suppose without loss of generality that $f : Y \to Y$ is the tower (5.1).

Lemma 5.8 *The family $\|W_n\|_{p\text{-var}}$ is tight on (Z, μ_Z).*

Proof Let $\tau_n = \sum_{j=0}^{n-1} \tau \circ F^j$ and define $U_n(t) = b_n^{-1} \sum_{j=0}^{\lfloor \tau_n t \rfloor - 1} v \circ f^j$ on Z. Note that $\|W_n\|_{p\text{-var}} \le \|U_n\|_{p\text{-var}}$. Let $s_i = \tau_i/\tau_n$, $i = 0, \dots, n$ and write $U_n = U_n' + U_n''$ where $U_n'|_{[s_i, s_{i+1})} \equiv U_n(s_i)$.

Observe that U_n' is a time-changed version of $\widetilde{W}_n$ (indeed $U_n'(s_i) = \widetilde{W}_n(i/n)$), so $\|U_n'\|_{p\text{-var}} = \|\widetilde{W}_n\|_{p\text{-var}}$. Thus the family $\|U_n'\|_{p\text{-var}}$ is tight on (Z, μ_Z).

Further we bound $\int_Z \|U_n''\|_{p\text{-var}} d\mu_Z$. Note that $U_n''(s_i) = 0$ and $\|1_{[s_i, s_{i+1})} U_n''\|_\infty \le b_n^{-1} \|v\|_\infty \tau \circ F^i$. Hence for $t \in [s_i, s_{i+1})$, $t' \in [s_{i'}, s_{i'+1})$,

$$\begin{aligned} |U_n(t) - U_n(t')|^p &\le \big(b_n^{-1} \|v\|_\infty \big(\tau \circ F^i + \tau \circ F^{i'}\big)\big)^p \\ &\le 2^{p-1} b_n^{-p} \|v\|_\infty^p \big(\tau^p \circ F^i + \tau^p \circ F^{i'}\big). \end{aligned}$$

It follows that

$$\|U_n''\|_{p\text{-var}}^p \le \sum_{i=0}^{n-1} \|U_n''\|_{p\text{-var}, [s_i, s_{i+1}]}^p + 2^p b_n^{-p} \|v\|_\infty^p \sum_{i=0}^{n-1} \tau^p \circ F^i.$$

On $[s_i, s_{i+1}]$, there are $\tau \circ F^i - 1$ jumps of size at most $b_n^{-1}\|v\|_\infty$, and one jump of size at most $b_n^{-1}\|v\|_\infty \tau \circ F^i$, so $\|U_n''\|_{p\text{-var}, [s_i, s_{i+1}]} \le \|U_n''\|_{1\text{-var}, [s_i, s_{i+1}]} \le 2b_n^{-1} \|v\|_\infty \tau \circ F^i$. Altogether, we have shown that

$$\|U_n''\|_{p\text{-var}} \lesssim \|v\|_\infty b_n^{-1} \left(\sum_{j=0}^{n-1} \tau^p \circ F^j \right)^{1/p}.$$

Now apply Proposition 4.9(c). □

Lemma 5.9 *The family $\|W_n\|_{p\text{-var}}$ is tight on (Y, μ_Z) if and only if it is tight on (Y, μ).*

Proof Observe that $W_n(t) \circ f = W_n(t + \frac{1}{n}) - b_n^{-1} v$ for all $t \ge 0$. Hence

$$\big| \|W_n\|_{p\text{-var}} - \|W_n\|_{p\text{-var}} \circ f \big| \le b_n^{-1}(|v| + |v| \circ f^n) \to_\mu 0.$$

Hence by Zweimüller [47, Theorem 1], $\|W_{n_k}\|_{p\text{-var}}$ has the same limit in distribution (if any) on (Y, μ_Z) as on (Y, μ) for each subsequence n_k. The result follows. □

Proof of Theorem 5.2 Combine Lemmas 5.8 and 5.9. □

6 Results for nonuniformly expanding maps

In this section, we prove results on weak convergence to a Lévy process, and tightness in p-variation, for a class of nonuniformly expanding maps. The weak convergence result extends work of [33] from scalar-valued observables to $\mathbb{R}^d$-valued observables. The result on tightness in p-variation is again new even for $d = 1$.

We show that intermittent maps such as (1.4) and (1.5) fit our setting in Sect. 6.2.

6.1 Nonuniformly expanding maps

Let $f: Y \to Y$ be a measurable transformation on a bounded metric space (Y, d) and let ν be a finite Borel measure on Y. Suppose that there exists a Borel subset $Z \subset Y$ with $\nu(Z) > 0$ and an at most countable partition $\mathcal{P}$ of Z (up to a zero measure set) with $\nu(a) > 0$ for each $a \in \mathcal{P}$. Suppose also that there is an integrable *return time* function $\tau: Z \to \{1, 2, \dots\}$ which is constant on each $a \in \mathcal{P}$ with value $\tau(a)$, such that $f^{\tau(a)}(z) \in Z$ for all $z \in a$, $a \in \mathcal{P}$.

Define the *induced map* $F: Z \to Z$, $F(z) = f^{\tau(z)}(z)$. We assume that f is *nonuniformly expanding*. That is, F is Gibbs–Markov as in Sect. 4 and in addition there is a constant $C > 0$ such that

$$d\big(f^k z, f^k z'\big) \le C d(Fz, Fz') \qquad \text{for all } 0 \le k \le \tau(a), z, z' \in a, a \in \mathcal{P}. \tag{6.1}$$

Let μ_Z be the unique F-invariant probability measure absolutely continuous with respect to ν. Define the ergodic f-invariant probability measure $\mu = \pi_* \mu_\Delta$ as in Sect. 5. Set $\bar\tau = \int_Z \tau \, \mathrm{d}\mu_Z$.

Let $v: Y \to \mathbb{R}^d$ be a Hölder observable with $\int_Y v \mathrm{d}\mu = 0$, and define V, $V^*: Z \to \mathbb{R}^d$ as in (5.2) and (5.3).

Let b_n be a sequence of positive numbers and define W_n as in (5.4). Let $\mathbb{P}$ be any probability measure on Y that is absolutely continuous with respect to ν, and regard W_n as a process with paths in $D([0, 1], \mathbb{R}^d)$, defined on the probability space $(Y, \mathbb{P})$.

We can now state and prove the main results of this subsection.

Theorem 6.1 *Suppose that:*

(a) $V: Z \to \mathbb{R}^d$ *is regularly varying on* (Z, μ_Z) *with index* $\alpha \in (1, 2)$ *and* σ *as in Definition* 4.1.
(b) b_n *satisfies* $\lim_{n\to\infty} n\mu_Z(|V| > b_n) = 1$.
(c) $V - \mathbb{E}(V \mid \mathcal{P}) \in L^p$ *for some* $p > \alpha$*, where* $\mathbb{E}$ *denotes the expectation on* (Z, μ_Z).
(d) $b_n^{-1} \max_{k<n} V^* \circ F^k \to_w 0$ *on* (Z, μ_Z).

Then $W_n \to_w L_\alpha$ *on* $(Y, \mathbb{P})$ *in the* $\mathcal{SM}_1$ *topology, where* L_α *is the* α*-stable Lévy process with spectral measure* $\Lambda = \cos\frac{\pi\alpha}{2}\Gamma(1-\alpha)\sigma/\bar\tau$.

Proof Note that $|V| \le \|v\|_\infty \tau$. Let $z, z' \in a$, $a \in \mathcal{P}$. Then

$$|V(z) - V(z')| \le \sum_{j=0}^{\tau(z)-1} |v(f^j z) - v(f^j z')| \le C_0 \sum_{j=0}^{\tau(z)-1} d(f^j z, f^j z')^\theta \le C_0 \tau(a) d(Fz, Fz')^\theta,$$

where C_0 is the Hölder constant for v and θ is the Hölder exponent, and we used condition (6.1) in the definition of nonuniformly expanding map. Hence condition (4.2) is satisfied.

Define $\widetilde{W}_n$ as in (5.4). By Theorem 4.2, $\widetilde{W}_n \to_w \tilde{L}_\alpha$ on (Z, μ_Z) in the $\mathcal{SJ}_1$ topology where $\tilde{L}_\alpha$ is an α-stable Lévy process with $\tilde{L}_\alpha$ having spectral measure $\tilde{\Lambda} = \cos\frac{\pi\alpha}{2}\Gamma(1-\alpha)\sigma$.

By Theorem 5.1, $W_n \to_w L_\alpha$ on (Y, μ) in the $\mathcal{SM}_1$ topology where $L_\alpha(t) = \tilde{L}_\alpha(t/\bar{\tau})$. This proves the result when $\mathbb{P} = \mu$.

By Zweimüller [47, Theorem 1 and Corollary 3] (see also [33, Proposition 2.8]), the convergence holds not only on (Y, μ) but also on $(Y, \mathbb{P})$ for any probability measure $\mathbb{P}$ that is absolutely continuous with respect to ν. This completes the proof. □

Theorem 6.2 *Suppose that τ is regularly varying with index $\alpha > 1$ on (Z, μ_Z), and that b_n satisfies $\lim_{n\to\infty} n\mu_Z(\tau > b_n) = 1$. Then $\{\|W_n\|_{p\text{-var}}\}$ is tight on $(Y, \mathbb{P})$ for each $p > \alpha$.*

Proof Condition (4.2) was established in the proof of Theorem 6.1. Tightness on (Y, μ) follows from Theorems 5.2 and 4.4. Tightness on $(Y, \mathbb{P})$ holds by the same argument used in the proof of Lemma 5.9. □

6.2 Intermittent maps

In this subsection, we show that Theorems 1.1 and 1.3 hold for the intermittent maps $f : [0, 1] \to [0, 1]$, given by (1.4) and (1.5).

We choose $Z = [\frac{1}{2}, 1]$ for the map (1.4), and $Z = [\frac{1}{3}, \frac{2}{3}]$ for (1.5). Let τ be the first return time to Z. The reference measure ν is Lebesgue and the partition $\mathcal{P}$ consists of maximal intervals on which the return time is constant. It is standard that the first return map $F = f^\tau$ is Gibbs–Markov, and since $f' > 1$, condition (6.1) holds. Thus both maps are nonuniformly expanding.

Lemma 6.3 *Let $v : [0, 1] \to \mathbb{R}^d$ be Hölder with $\int v\,d\mu = 0$ and $v(0) \ne 0$, also $v(1) \ne 0$ in case f is given by (1.5). Define V, $V^* : Z \to \mathbb{R}^d$ as in (5.2) and (5.3). Then*

(a) *There exists a unique absolutely continuous f-invariant probability measure μ on $[0, 1]$. Its density h is bounded below and is continuous on Z.*
(b) *V is regularly varying with index α on (Z, μ_Z). The probability measure σ as in Definition 4.1 is given by*

$$\sigma = \begin{cases} \delta_{v(0)/|v(0)|} & \text{for the map (1.4)}, \\ \frac{|v(0)|^\alpha}{|v(0)|^\alpha + |v(1)|^\alpha}\delta_{v(0)/|v(0)|} + \frac{|v(1)|^\alpha}{|v(0)|^\alpha + |v(1)|^\alpha}\delta_{v(1)/|v(1)|} & \text{for the map (1.5)}. \end{cases}$$

(c) $\lim_{n\to\infty} n\mu_Z(|V| > b_n) = 1$ *with* $b_n = c^{1/\alpha} n^{1/\alpha}$, *where*

$$c = \begin{cases} \frac{1}{4}|v(0)|^\alpha \alpha^\alpha h(\frac{1}{2})\bar{\tau} & \text{for the map (1.4)}, \\ \frac{1}{9}\big(|v(0)|^\alpha + |v(1)|^\alpha\big)\alpha^\alpha h(\frac{1}{3})\bar{\tau} & \text{for the map (1.5)}. \end{cases}$$

Here $\bar{\tau} = \int_Z \tau \, d\mu_Z$.

(d) $V - \mathbb{E}(V \mid \mathcal{P}) \in L^p$ *for some* $p > \alpha$.

(e) $n^{-1/\alpha} \max_{0\le k<n} V^* \circ F^k \to_w 0$ *on* (Z, μ_Z).

Proof We give the details for the map (1.5). The details for the map (1.4) are similar and simpler.

Let $a_1 = \frac{1}{3}$ and $a_k = a_{k+1}(1 + (3a_{k+1})^{1/\alpha})$, $k \ge 1$. By a standard calculation, see for example [19], $a_k \sim \frac{1}{3}\alpha^\alpha k^{-\alpha}$. Let $z_k = \frac{1}{3}(a_k + 1)$ and $z'_k = 1 - z_k$. The partition $\mathcal{P}$ consists of the intervals (z_k, z_{k-1}) and (z'_{k-1}, z'_k), $k \ge 2$, on which τ equals k, and (z_1, z'_1) where τ equals 1.

Observe that $F = f^\tau$ has full branches, i.e. $Fa = Z$ for every $a \in \mathcal{P}$, modulo zero measure. It is standard that the unique F-invariant absolutely continuous measure μ_Z has continuous density h_Z bounded away from zero (see for example [23, Proposition 2.5]). Moreover, h is bounded below and $h|_Z = h_Z/\bar{\tau}$.

If $z \in (\frac{1}{3}, z_k)$ and $0 < \ell \le k$, then $f^\ell z \in (0, a_{k-\ell+1})$, so $|f^\ell z| \lesssim (k-\ell)^{-\alpha}$. Similarly, if $z \in (z'_k, \frac{2}{3})$, then $|1 - f^\ell z| \lesssim (k-\ell)^{-\alpha}$. Let $\theta \in (0, 1]$ be the Hölder exponent of v. Without loss, we assume that $\theta < 1/\alpha$. Define $\hat{v} = v(0)1_{(\frac{1}{3},\frac{1}{2})} + v(1)1_{(\frac{1}{2},\frac{2}{3})}$ on Z. Then

$$\left|\ell\hat{v}(z) - \sum_{j=0}^{\ell-1} v(f^j z)\right| \le |\hat{v}(z) - v(z)| + \sum_{j=1}^{\tau(z)-1} |\hat{v}(z) - v(f^j z)| \lesssim \tau(z)^\beta \tag{6.2}$$

for $\ell \le \tau(z)$, where $\beta = 1 - \alpha\theta \in (0, 1)$. In particular, $|\tau\hat{v} - V| \lesssim \tau^\beta$.

By symmetry and continuity of h_Z,

$$\mu_Z\big(z > \tfrac{1}{2},\ \tau > k\big) = \mu_Z(z < \tfrac{1}{2},\ \tau > k) = \mu_Z\big((\tfrac{1}{3}, z_k)\big) \sim \frac{h_Z\big(\frac{1}{3}\big)\alpha^\alpha}{9k^\alpha}.$$

Let B be a Borel set in $\mathbb{S}^{d-1}$ and suppose that $v(0)/|v(0)| \in B$, $v(1)/|v(1)| \notin B$. Then

$$\begin{aligned} \frac{\mu_Z(|\tau\hat{v}| > rt,\ \tau\hat{v}/|\tau\hat{v}| \in B)}{\mu_Z(|\tau\hat{v}| > t)} &= \frac{\mu_Z(z < \frac{1}{2},\ \tau > rt/|v(0)|)}{\mu_Z(z < \frac{1}{2},\ \tau > t/|v(0)|) + \mu_Z(z > \frac{1}{2},\ \tau > t/|v(1)|)} \\ &\to r^{-\alpha}\frac{|v(0)|^\alpha}{|v(0)|^\alpha + |v(1)|^\alpha} \quad \text{as } t \to \infty. \end{aligned}$$

The calculations for the remaining Borel sets B are similar, and it follows that $\tau\hat{v}$ is regularly varying with index α and that the probability measure σ as in Definition 4.1

is given by the formula in part (b). By (6.2), V is regularly varying with index α and the same σ, proving part (b).

Moreover, $\mu_Z(|\tau\hat{v}| > n) \sim cn^{-\alpha}$ with c as in part (c), so $\mu_Z(|V| > n) \sim cn^{-\alpha}$ by (6.2). Part (c) follows by Remark 4.3(a).

It is immediate from (6.2) that $|V(z) - V(z')| \lesssim \tau(a)^\beta$ for all $z, z' \in a$, $a \in \mathcal{P}$. Part (d) follows by Remark 4.3(b).

Finally, it follows from (6.2) that $V^* \lesssim \tau^\beta$, from which $V^* \in L^q(\mu_Z)$ for some $q > \alpha$, and

$$\int \Big(n^{-1/\alpha} \max_{0\le k<n} V^* \circ F^k\Big)^q d\mu_Z \le n^{-q/\alpha} \sum_{k<n} \int (V^*)^q \circ F^k d\mu_Z$$
$$= n^{-q/\alpha+1} \|V^*\|_q^q \to 0.$$

This proves (e) and completes the proof of the lemma. □

Theorems 1.1 and 1.3 now follow from Theorems 6.1 and 6.2. Moreover, L_α is identified as the α-stable Lévy process with spectral measure $\Lambda = c\cos\frac{\pi\alpha}{2}\Gamma(1-\alpha)\sigma/\bar{\tau}$ with c and σ as in Lemma 6.3.

Finally, as a consequence of these results combined with Theorem 2.6, we can record the desired conclusion for homogenisation of fast–slow systems with fast dynamics given by one of the intermittent maps in Sect. 1.

Corollary 6.4 *Consider the intermittent map (1.4) or (1.5) with $\alpha \in (1, 2)$ and let $v\colon Y \to \mathbb{R}^d$ be Hölder with $\int_Y v d\mu = 0$ and $v(0) \neq 0$, also $v(1) \neq 0$ in case of (1.5).*

Consider the fast–slow system (1.1) with initial condition $x_0^{(n)} = \xi_n$ such that $\lim_{n\to\infty} \xi_n = \xi$. Suppose that $a \in C^\beta(\mathbb{R}^m, \mathbb{R}^m)$, $b \in C^\gamma(\mathbb{R}^m, \mathbb{R}^{m\times d})$ for some $\beta > 1$, $\gamma > \alpha$. Define W_n as in (1.2) and $X_n(t) = x^{(n)}_{\lfloor nt\rfloor}$. Let $\mathbb{P}$ be any probability measure on Y that is absolutely continuous with respect to Lebesgue, and regard W_n and X_n as processes on $(Y, \mathbb{P})$.

Let ℓ_k denote the linear path function on $\mathbb{R}^k$ and let ϕ_b be the path function on $\mathbb{R}^{d+m}$ as in Definition 2.5. Fix $p > \alpha$. Then

$$((W_n, X_n), \ell_{d+m}) \to_w ((L_\alpha, X), \phi_b) \quad \text{as} \quad n \to \infty$$

in $(\mathscr{D}^{p\text{-var}}([0, 1], \mathbb{R}^{d+m}), \boldsymbol{\alpha}_{p\text{-var}})$, where L_α is the α-stable Lévy process with spectral measure $\Lambda = c\cos\frac{\pi\alpha}{2}\Gamma(1-\alpha)\sigma/\bar{\tau}$ with c and σ as in Lemma 6.3, and X is the solution of the Marcus differential equation (2.4). □

Acknowledgements I.C. was funded by a Junior Research Fellowship of St John's College, Oxford while this work was carried out. P.K.F. acknowledges partial support from the ERC, CoG-683164, the Einstein Foundation Berlin, and DFG research unit FOR2402. A.K. and I.M. acknowledge partial support from the European Advanced Grant StochExtHomog (ERC AdG 320977). A.K. is also supported by an Engineering and Physical Sciences Research Council Grant EP/P034489/1. We would like to thank the anonymous referees for their helpful and detailed comments.

References

1. Aaronson, J., Denker, M.: Local limit theorems for partial sums of stationary sequences generated by Gibbs–Markov maps. Stoch. Dyn. **1**, 193–237 (2001)
2. Applebaum, D.: Lévy Processes and Stochastic Calculus, vol. 116, 2nd edn. Cambridge Stud. Adv. Math, Cambridge (2009)
3. Bingham, N.H., Goldie, C.M., Teugels, J.L.: Regular Variation. Encyclopedia of Mathematics and its Applications, vol. 27. Cambridge University Press, Cambridge (1987)
4. Bradley, R.C.: Basic properties of strong mixing conditions a survey and some open questions. Probab. Surv. **2**, 107–144 (2005)
5. Chechkin, A., Pavlyukevich, I.: Marcus versus Stratonovich for systems with jump noise. J. Phys. A **47**, 342001 (2014)
6. Chevyrev, I.: Random walks and Lévy processes as rough paths. Probab. Theory Relat. Fields **170**, 891–932 (2018)
7. Chevyrev, I., Friz, P.K.: Canonical RDEs and general semimartingales as rough paths. Ann. Probab. **47**, 420–463 (2019)
8. Chevyrev, I., Friz, P.K., Korepanov, A., Melbourne, I., Zhang, H.: Multiscale systems, homogenization, and rough paths. In: Friz, P., et al. (eds.) Probability and Analysis in Interacting Physical Systems: In Honor of S.R.S. Varadhan, Berlin, August, 2016 (.), Springer Proceedings in Mathematics and Statistics, vol. 283, pp. 17–42 (2019)
9. Chevyrev, I., Friz, P.K., Korepanov, A., Melbourne, I., Zhang, H.: Deterministic homogenization under optimal moment assumptions for fast–slow systems. Part 2. Preprint (2019)
10. Dolgopyat, D.: Limit theorems for partially hyperbolic systems. Trans. Am. Math. Soc. **356**, 1637–1689 (2004)
11. Dolgopyat, D.: Averaging and invariant measures. Mosc. Math. J. **5**, 537–576 (2005)
12. Friz, P., Hairer, M.: A Course on Rough Paths. With an Introduction to Regularity Structures. Universitext. Springer, Berlin (2014)
13. Friz, P.K., Shekhar, A.: General rough integration, Lévy rough paths and a Lévy–Kintchine-type formula. Ann. Probab. **45**, 2707–2765 (2017)
14. Friz, P.K., Victoir, N.B.: Multidimensional Stochastic Processes as Rough Paths, vol. 120. Cambridge Stud. Adv. Math, Cambridge (2010)
15. Friz, P.K., Zhang, H.: Differential equations driven by rough paths with jumps. J. Differ. Equ. **264**, 6226–6301 (2018)
16. Gottwald, G., Melbourne, I.: Homogenization for deterministic maps and multiplicative noise. Proc. R. Soc. Lond. A **469**, 20130201 (2013)
17. Gouëzel, S.: Central limit theorem and stable laws for intermittent maps. Probab. Theory Relat. Fields **128**, 82–122 (2004)
18. Gouëzel, S.: Statistical properties of a skew product with a curve of neutral points. Ergod. Theory Dyn. Syst. **27**, 123–151 (2007)
19. Holland, M.: Slowly mixing systems and intermittency maps. Ergod. Theory Dyn. Syst. **25**, 133–159 (2005)
20. Kelly, D., Melbourne, I.: Smooth approximation of stochastic differential equations. Ann. Probab. **44**, 479–520 (2016)
21. Kelly, D., Melbourne, I.: Homogenization for deterministic fast–slow systems with multidimensional multiplicative noise. J. Funct. Anal. **272**, 4063–4102 (2017)
22. Kocheim, D., Pühringer, F., Zweimüller, R.: A functional stable limit theorem for Gibbs–Markov maps. Preprint (2018)
23. Korepanov, A., Kosloff, Z., Melbourne, I.: Explicit coupling argument for nonuniformly hyperbolic transformations. Proc. Edinb. Math. Soc. **149**, 101–130 (2019)

24. Korepanov, A., Kosloff, Z., Melbourne, I.: Deterministic homogenization under optimal moment assumptions for fast–slow systems. Part 1. Preprint (2020)
25. Kurtz, T.G., Pardoux, E., Protter, P.: Stratonovich stochastic differential equations driven by general semimartingales. Ann. Inst. H. Poincaré Probab. Stat. **31**, 351–377 (1995)
26. Lépingle, D.: La variation d'ordre p des semi-martingales. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **36**, 295–316 (1976)
27. Liverani, C., Saussol, B., Vaienti, S.: A probabilistic approach to intermittency. Ergod. Theory Dyn. Syst. **19**, 671–685 (1999)
28. Lyons, T.: Differential equations driven by rough signals. I. An extension of an inequality of L.C. Young. Math. Res. Lett. **1**, 451–464 (1994)
29. Marcus, S.I.: Modeling and approximation of stochastic differential equations driven by semimartingales. Stochastics **4**, 223–245 (1980/81)
30. Melbourne, I., Nicol, M.: Almost sure invariance principle for nonuniformly hyperbolic systems. Commun. Math. Phys. **260**, 131–146 (2005)
31. Melbourne, I., Török, A.: Statistical limit theorems for suspension flows. Isr. J. Math. **144**, 191–209 (2004)
32. Melbourne, I., Stuart, A.: A note on diffusion limits of chaotic skew product flows. Nonlinearity **24**, 1361–1367 (2011)
33. Melbourne, I., Zweimüller, R.: Weak convergence to stable Lévy processes for nonuniformly hyperbolic dynamical systems. Ann Inst. H. Poincaré (B) Probab. Stat. **51**, 545–556 (2015)
34. Pavliotis, G.A., Stuart, A.M.: Multiscale Methods. Texts in Applied Mathematics, vol. 53. Springer, New York (2008). Averaging and homogenization
35. Pène, F.: Averaging method for differential equations perturbed by dynamical systems. ESAIM Probab. Stat. **6**, 33–88 (2002)
36. Pisier, G., Xu, Q.H.: The strong p-variation of martingales and orthogonal series. Probab. Theory Relat. Fields **77**, 497–514 (1988)
37. Pomeau, Y., Manneville, P.: Intermittent transition to turbulence in dissipative dynamical systems. Commun. Math. Phys. **74**, 189–197 (1980)
38. Ratner, M.: The central limit theorem for geodesic flows on n-dimensional manifolds of negative curvature. Isr. J. Math. **16**, 181–197 (1973)
39. Samorodnitsky, G., Taqqu, M.: Stable non-Gaussian random processes: stochastic models with infinite variance. Chapman and Hall, London (1994)
40. Skorohod, A.V.: Limit theorems for stochastic processes. Theory Probab. Appl. **1**, 261–290 (1956)
41. Tyran-Kamińska, M.: Convergence to Lévy stable processes under some weak dependence conditions. Stochastic Process. Appl. **120**, 1629–1650 (2010)
42. Tyran-Kamińska, M.: Weak convergence to Lévy stable processes in dynamical systems. Stoch. Dyn. **10**, 263–289 (2010)
43. Whitt, W.: Stochastic-process limits. Springer-Verlag, New York (2002)
44. Williams, D.: Path-wise solutions of stochastic differential equations driven by Lévy processes. Rev. Mat. Iberoam **17**, 295–329 (2001)
45. Wong, E., Zakai, M.: On the convergence of ordinary integrals to stochastic integrals. Ann. Math. Statist. **36**, 1560–1564 (1965)
46. Zweimüller, R.: Stable limits for probability preserving maps with indifferent fixed points. Stoch. Dyn. **3**, 83–99 (2003)
47. Zweimüller, R.: Mixing limit theorems for ergodic transformations. J. Theoret. Probab. **20**, 1059–1071 (2007)