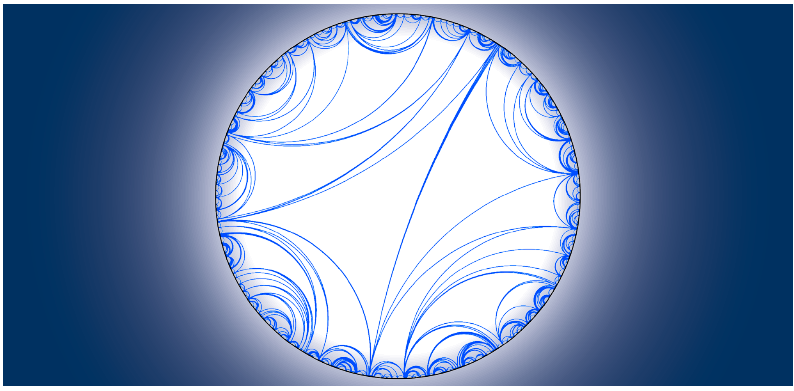
Universität Potsdam



Michael Högele | Paulo Ruffino

Averaging along Lévy diffusions in foliated spaces

Preprints des Instituts für Mathematik der Universität Potsdam
2 (2013) 10

Preprints des Instituts für Mathematik der Universität Potsdam

Preprints des Instituts für Mathematik der Universität Potsdam
2 (2013) 10

Michael Högele | Paulo Ruffino

Averaging along Lévy diffusions in foliated spaces

Universitätsverlag Potsdam

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.de abrufbar.

Universitätsverlag Potsdam 2013
http://info.ub.uni-potsdam.de/verlag.htm

Am Neuen Palais 10, 14469 Potsdam
Tel.: +49 (0)331 977 2533 / Fax: 2292
E-Mail: verlag@uni-potsdam.de

Die Schriftenreihe **Preprints des Instituts für Mathematik der Universität Potsdam** wird herausgegeben vom Institut für Mathematik der Universität Potsdam.

ISSN (online) 2193-6943

Kontakt:
Institut für Mathematik
Am Neuen Palais 10
14469 Potsdam
Tel.: +49 (0)331 977 1028
WWW: http://www.math.uni-potsdam.de

Titelabbildungen:
1. Karla Fritze | Institutsgebäude auf dem Campus Neues Palais
2. Nicolas Curien, Wendelin Werner | Random hyperbolic triangulation
Published at: http://arxiv.org/abs/1105.5089

Online veröffentlicht auf dem Publikationsserver der Universität Potsdam
URL http://pub.ub.uni-potsdam.de/volltexte/2013/6492/
URN urn:nbn:de:kobv:517-opus-64926
http://nbn-resolving.de/urn:nbn:de:kobv:517-opus-64926

Averaging along Lévy diffusions in foliated spaces

Michael Högele* Paulo Ruffino†

April 13, 2013‡

Abstract

We consider an SDE driven by a Lévy noise on a foliated manifold, whose trajectories stay on compact leaves. We determine the effective behavior of the system subject to a small smooth transversal perturbation of order $\varepsilon > 0$. More precisely, we show that the average of the transversal component of the SDE converges to the solution of a deterministic ODE, according to the average of the perturbing vector field with respect to the invariant measures on the leaves (of the unpertubed system) as ε goes to 0. In particular we give upper bounds for the rates of convergence. The main results which are proved for pure jump Lévy processes complement the result by Gargate and Ruffino for Stratonovich SDEs to Lévy driven SDEs of Marcus type.

Keywords: averaging principle; foliated diffusion; Lévy diffusions on manifolds; canonical Marcus integration

2010 Mathematical Subject Classification: 60H10, 60J60, 60G51, 58J65, 58J37.

1 Introduction

The purpose of this article is to extend the scope of an averaging principle from continuous semimartingales in Gonzáles and Ruffino [4] to Lévy diffusions containing a pure jump component.

The system under consideration is the following. We study a stochastic differential equation (SDE) driven by a jump Lévy noise with values in a smooth Riemannian manifold M with a foliation structure $\mathfrak{M}$, i.e. there exists an equivalence relation on M, which defines a family of immersed submanifolds (the elements of $\mathfrak{M}$, called the leaves of the foliation) of constant dimension r. For more details and further properties on foliated spaces we refer to Tondeur [12] or Walcak [13]. The solution flow of the SDE is assumed to be foliated with respect to $\mathfrak{M}$ in the sense that each of the (discontinuous) solution paths of the SDE stays on the corresponding leaf of its initial condition almost surely for all times. We further assume the existence of a unique invariant measure for the SDE on each leaf.

If this system is perturbed by a smooth deterministic vector field εK transversal to the leaves with intensity $\varepsilon > 0$, the foliated structure of the solution is destroyed due to the appearance of a (smooth) transversal component in the trajectories. We study the effective behavior of this transversal component in the limit as ε tends to 0.

The main idea is the following. Consider the solution along the rescaled time t/ε, its foliated component approximates the ergodic average behavior for small ε. Hence the essential transversal

*Institut für Mathematik, Universität Potsdam, Germany; hoegele@uni-potsdam.de

†IMECC, Universidade Estadual de Campinas, Brazil; ruffino@ime.unicamp.br

‡Averaging8.tex

behavior is captured by an ODE for the transversal component driven by the vector field K instead of εK, which is averaged by the ergodic invariant measure on the leaves. Note that the intensity of the original pertubation εK cancels out by the time scaling t/ε. This is the result of Theorem 4.1 and will be referred to as an averaging principle. Our calculations here also determine upper bounds for the rates of convergence and a probabilistic robustness result.

The heuristics of an averaging principle consists in replacing the fine dynamical impact of a so-called fast variable on the dynamics of a so-called slow variable by its averaged statistical static influence. For references on the vast also classical literature on averaging for deterministic systems see e.g. the books by V. Arnold [2] and Saunders, Verhulst and Murdock [11] and the numerous citations therein. For stochastic systems among many others we mention the book by Kabanov and Pergamenshchikov [5] and the references therein which gives an excellent overview on the subject. See also [3, 6]. An inspiration for this article also goes back to the work [8] by Li, where she established an averaging principle for the particular case of completely integrable (continuous) stochastic Hamiltonian systems. In Gonzáles and Ruffino [4] these results have been generalized to averaging principles for perturbations of Gaussian diffusions on foliated spaces. This article completes this result for general Lévy driven foliated diffusions.

The article is organized as follows. In the next paragraph we present the precise framework of our study. Section 2 is dedicated to the fundamental technical Proposition 2.1, where the stochastic Marcus integral technique is applied and whose estimates turn out to be the basis for the rates of convergence of the main theorem. Section 3 deals with the averaging on the leaves. In Section 4 we prove the main theorem and provide a simple, but instructive example.

The set up. Let M be a simply connected smooth Riemannian manifold with an n-dimensional smooth foliation. We denote by L_x the leaf of the foliation passing through a point $x \in M$. For simplicity we assume that the leaves are compact an that each leaf L_x has a tubular neighborhood $U \subset M$, where U is diffeomorphic to $L_x \times V$ where $V \subset \mathbb{R}^d$ is an open bounded neighborhood of the origin and d is the codimension of the foliation.

We shall assume an SDE in M immersed in an Euclidean space, whose solution preserves the foliation, i.e. it is a Marcus canonical equation also known as generalized Stratonovich equation in the sense of Kurtz, Pardoux and Protter [7]. We consider the foliated stochastic differential equation

$$dX_t = F_0(X_t)dt + F(X_t) \diamond dZ_t, \qquad X_0 = x_0, \tag{1}$$

which consists of the following components.

1. Let $F \in \mathcal{C}^1(M; L(\mathbb{R}^r; T\mathfrak{M}))$, such that the map $x \mapsto F(x)$ is $\mathcal{C}^1$ and the linear map $F(x) : \mathbb{R}^r \to T_x L_x$. We write $F_i(x) = F(x)e_i$, for e_i the canonical basis of $\mathbb{R}^r$.

2. Here $Z = (Z_t)_{t \geqslant 0}$ with $Z_t = (Z_t^1, \dots, Z_t^r)$ is a Lévy process in $\mathbb{R}^r$ with respect to a given filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geqslant 0}, \mathbb{P})$ with characteristic triplet $(0, \nu, 0)$. It is assumed that the filtration satisfies the usual conditions in the sense of Protter [9]. It is a consequence of the Lévy-Itô decomposition of Z (see for instance Applebaum [1]) that Z is a pure jump process with respect to a Lévy measure $\nu : \mathcal{B}(\mathbb{R}^r) \to [0, \infty]$ which satisfies $\int (1 \wedge \|y\|^2) \nu(dy) < \infty$ of the pure jump part Z^d of Z and $b \in \mathbb{R}^d$ the linear drift component of Z. Due to the compactness of the leaves we restrict ourselves to the case of uniformly bounded jumps of Z, i.e.

$$\exists\, R > 0 : \quad \nu(B_R^c(0)) = 0.$$

3. Equation (1) is defined as

$$X_t^\varepsilon = x_0 + \int_0^t F_0(X_s^\varepsilon)ds + \int_0^t F(X_{s-}^\varepsilon)dZ_s + \sum_{0<s\leqslant t} (\Phi^{F\Delta_s Z}(X_{s-}^\varepsilon) - X_{s-}^\varepsilon) - F(X_{s-}^\varepsilon)\Delta_s Z) \quad (2)$$

where $\Phi^{Fz}(x) = Y(1, x; Fz)$, where $Y(t, x; Fz)$ is the solution of the ordinary differential equation with initial condition

$$\frac{d}{d\sigma} Y(\sigma) = F(Y(\sigma))z \qquad Y(0) = x \in M, \quad z \in \mathbb{R}^r.$$

4. The upper bound of the jump size is such that, that there exists a compromise between the vector field and the domain of a local coordinate system. Precisely, for given upper bound $R > 0$ for the increments we assume the vector field F to satisfy that for each point $x \in M$ there exist local coordinates φ such that $x, \Phi^{Fz}(x) \in \mathrm{dom}(\varphi)$ for all $z \in B_R(0) \subset \mathbb{R}^r$.

   Point 4 yields that if $\Delta_s Z \neq 0$ and $X_{s-}(x_0) \in \mathrm{dom}(\varphi)$, then also $X_s(x_0) = X_{s-}(x_0) + \Delta_s X(x_0)$ belongs to $\mathrm{dom}(\varphi)$. It can always be satisfied, since $\Delta_s X$ depends continuously on $F(X_{s-})\Delta_s Z$ and the radius of injectivity of the compact leaf, which represents the overall worst case on the manifold, is always positive. This is a moderate technical restriction, which does not disturb the spirit of the result, since the leaves of M are compact and the jumps of Z are uniformly bounded. It is necessary in order to avoid implicit conditioning on the nonappearance of jumps beyond a certain size.

Proposition 4.3 in Kurtz, Pardoux and Protter [7] shows a sort of support theorem, i.e. that under the aforementioned conditions $\mathbb{P}(X_0 \in L_{x_0}) = 1$ implies $\mathbb{P}(X_t \in L_{x_0} \quad \forall t \geqslant 0) = 1$. For a smooth vector field K in M, we shall denote by X^ε, $\varepsilon > 0$ the solution of the perturbed system

$$dX_t^\varepsilon = F_0(X_t^\varepsilon)dt + F(X_t) \diamond dZ_t + \varepsilon K(X_t^\varepsilon)dt, \qquad X_0^\varepsilon = x_0. \quad (3)$$

It is defined analoguously to equation (2).

2 Preliminary results

The local coordinates. Given an initial condition $x_0 \in M$ let $U \subset M$ be a bounded neighborhood of x_0 which is diffeomorphic to $L_{x_0} \times V$ whose closure $\bar{U} \subset M$. By compactness of L_{x_0}, there is a finite number k, say, of local foliated coordinate systems $\varphi_i : U_i \to W_i \times V \subset \mathbb{R}^n \times \mathbb{R}^d$ and $x_0 \in U_1$ with the following properties.

1. $U = \bigcup_{j=1}^k U_j$
2. $L_{x_0} = \bigcup_{j=1}^k \varphi_i^{-1}(W_i \times \{0\})$
3. For all $q_1 \in U_i$ and $q_2 \in U_j$, $i, j \in \{1, \ldots k\}$ we have that $L_{q_1} = L_{q_2}$ if the projection onto the transversal components, which have values in V is identical, that is $\pi_2(\varphi_i(q_1)) = \pi_2(\varphi_j(q_2))$.
4. The coordinates φ_i, $i = 1, \ldots, k$ have bounded derivatives.

Proposition 2.1 *Let τ^ε be the first time the solution $X^\varepsilon(x_0)$ of (3) exits the aforementioned foliated coordinate neighborhood U_1. Then for all Lipschitz test functions $\Psi : M \to \mathbb{R}$, exponent $p \in [2, \infty)$ there exist $K_1, K_2 > 0$ such that for all $T \geqslant 0$ it follows that*

$$\left(\mathbb{E}\left[\sup_{s \leqslant T \wedge \tau^\varepsilon} |\Psi(X_s^\varepsilon(x_0)) - \Psi(X_s(x_0))|^p\right]\right)^{\frac{1}{p}} \leqslant K_1 \, \varepsilon \, T \exp\left(K_2 \, T^p\right). \quad (4)$$

The constants K_1, K_2 depend on the upper bounds of the norms for the perturbing vector field K, of the Lipschitz coefficient of Ψ and the derivatives of the vector fields $F_0, F_1 \ldots, F_r$ with respect to the coordinate system.

Proof: Change of coordinates: First we rewrite X^ε and X, the solutions of equation (1) and (3), in terms of the foliated coordinates φ_1

$$(u_t, v_t) := \varphi_1(X_t) \qquad \text{and} \qquad (u_t^\varepsilon, v_t^\varepsilon) := \varphi_1(X_t^\varepsilon).$$

Exploiting the regularity of Ψ and φ_1 we obtain

$$|\Psi(X_t^\varepsilon) - \Psi(X_t)| = |\Psi \circ \varphi_1^{-1}(u_t^\varepsilon, v_t^\varepsilon) - \Psi \circ \varphi_1^{-1}(u_t, v_t)| \leqslant C'(|u_t^\varepsilon - u_t| + |v_t^\varepsilon - v_t|) \tag{5}$$

for $C' := Lip(\psi) \sup_{y \in \bar{U}_1} |\varphi_1^{-1}(y)|$. Further we define

$$\begin{aligned} \mathfrak{F}_i &:= (D\varphi_1) \circ F_i \circ \varphi_1^{-1} \qquad \text{for } i = 0, \ldots, n, \\ \mathfrak{K} &:= (D\varphi_1) \circ K \circ \varphi_1^{-1}, \end{aligned}$$

which all together with their derivatives are bounded. We adopt the notation

$$\mathfrak{K} = (\mathfrak{K}_1, \mathfrak{K}_2), \qquad \mathfrak{K}_1 = (\mathfrak{K}_1^1, \ldots, \mathfrak{K}_1^n), \qquad \mathfrak{K}_2 = (\mathfrak{K}_2^1, \ldots, \mathfrak{K}_2^d).$$

The chain rule proved in [7] Theorem 4.2 yields for equation (3) the following form in φ_1 coordinates written for the different components

$$du_t^{\varepsilon,i} = \pi_1^i \mathfrak{F}_0(u_t^\varepsilon, v_t^\varepsilon)dt + \pi_1^i \mathfrak{F}(u_t^\varepsilon, v_t^\varepsilon) \diamond dZ_t + \varepsilon\, \mathfrak{K}_1^i(u_t^\varepsilon, v_t^\varepsilon)dt \qquad \text{for } i = 1, \ldots, n, \tag{6}$$

$$dv_t^{\varepsilon,j} = \varepsilon\, \mathfrak{K}_2^j(u_t^\varepsilon, v_t^\varepsilon)dt \qquad \text{for } j = 1, \ldots, d, \tag{7}$$

where $\mathfrak{F}_i(x) = \nabla\varphi_1(x) F_i \circ \varphi_1^{-1}(x)$, $i = 1, \ldots, n$ and π_1^i is the projection onto the i-th coordinate of $W_1 \subset \mathbb{R}^n$. Equation (5) yields the estimate

$$\sup_{s \leqslant t \wedge \tau^\varepsilon} |v_s^\varepsilon - v_s| \leqslant \varepsilon \sup_{s \leqslant t \wedge \tau^\varepsilon} \int_0^s |\mathfrak{K}_2(u_\sigma^\varepsilon, v_\sigma^\varepsilon)|\, d\sigma \leqslant C_0\, \varepsilon\, t \tag{8}$$

where $C_0 = \sup_{y \in U} |\mathfrak{K}_1(y)|$. Analoguously we estimate for further purpose in abuse of notation with a constant

$$\varepsilon \sup_{s \leqslant t \wedge \tau^\varepsilon} \int_0^s |\mathfrak{K}_1(u_\sigma^\varepsilon, v_\sigma^\varepsilon)|\, d\sigma \leqslant C_1\, \varepsilon\, t. \tag{9}$$

By equation (6) we have that for $s < \tau^\varepsilon$

$$\begin{aligned} u_s^\varepsilon - u_s = &\int_0^s (\mathfrak{F}_0(u_\sigma^\varepsilon, v_\sigma^\varepsilon) - \mathfrak{F}_0(u_\sigma, v_\sigma))d\sigma \\ &+ \int_0^s (\mathfrak{F}(u_\sigma^\varepsilon, v_\sigma^\varepsilon) - \mathfrak{F}(u_\sigma, v_\sigma)) \diamond dZ_\sigma \\ &+ \varepsilon \int_0^s \mathfrak{K}_1(u_\sigma^\varepsilon, v_\sigma^\varepsilon)d\sigma. \end{aligned}$$

This equality is defined as

$$u_t^\varepsilon - u_t = \int_0^t [\mathfrak{F}_0(u_s^\varepsilon, v_s^\varepsilon) - \mathfrak{F}_0(u_s, v_s)]ds$$

$$
\begin{aligned}
&+\int_0^t [\mathfrak{F}(u^\varepsilon_{s-},v^\varepsilon_{s-})-\mathfrak{F}(u_{s-},v_{s-})]dZ_s\\
&+\sum_{0<s\leqslant t}\big[(\Phi^{\mathfrak{F}\Delta_s Z}(u^\varepsilon_{s-},v^\varepsilon_{s-})-\Phi^{\mathfrak{F}\Delta_s Z}(u_{s-},v_{s-}))\\
&\qquad -(u^\varepsilon_{s-}-u_{s-})-(\mathfrak{F}(u^\varepsilon_{s-},v^\varepsilon_{s-})-\mathfrak{F}(u_{s-},v_{s-}))\Delta_s Z\big]\\
&+\varepsilon\int_0^t \mathfrak{K}_1(u^\varepsilon_\sigma,v^\varepsilon_\sigma)d\sigma.
\end{aligned}
$$

Since $p \geqslant 1$ this leads to

$$
\begin{aligned}
|u^\varepsilon_t-u_t|^p &\leqslant 4^{p-1}\Big|\int_0^t \mathfrak{F}_0(u^\varepsilon_s,v^\varepsilon_s)-\mathfrak{F}_0(u_s,v_s)ds\Big|^p+\ 4^{p-1}C_1^p\varepsilon^p t^p\\
&+4^{p-1}\Big|\int_0^t [\mathfrak{F}(u^\varepsilon_{s-},v^\varepsilon_{s-})-\mathfrak{F}(u^\varepsilon_{s-},v^\varepsilon_{s-})]dZ_s\Big|\\
&+4^{p-1}\Big|\sum_{0<s\leqslant t}\Phi^{\mathfrak{F}\Delta_s Z}(u^\varepsilon_{s-},v^\varepsilon_{s-})-\Phi^{\mathfrak{F}\Delta_s Z}(u_{s-},v_{s-})-(u^{\varepsilon,i}_{s-}-u^i_{s-})\\
&-(\mathfrak{F}(u^\varepsilon_{s-},v^\varepsilon_{s-})-\mathfrak{F}(u_{s-},v_{s-}))\Delta_s Z\Big|^p.
\end{aligned}\tag{10}
$$

We now estimate the terms of the right-hand side in (10). The first term on the right-hand side is estimated in a straight-forward manner with the help of equation (8) by

$$
\begin{aligned}
\Big|\int_0^t \mathfrak{F}_0(u^\varepsilon_s,v^\varepsilon_s)-\mathfrak{F}_0(u_s,v_s)ds\Big|^p &\leqslant \Big(\int_0^t C_2|(u^\varepsilon_s-u_s,v^\varepsilon_s-v_s)|ds\Big)^p\\
&\leqslant (C_2C_3)^p\Big(\int_0^t(|u^\varepsilon_s-u_s|+|v^\varepsilon_s-v_s|)ds\Big)^p\\
&\leqslant (C_2C_3)^p t^{p-1}\left(\int_0^t |u^\varepsilon_s-u_s|^p ds+\int_0^t |v^\varepsilon_s-v_s|^p ds\right)\\
&\leqslant (C_2C_3)^p t^{p-1}\left(\int_0^t |u^\varepsilon_s-u_s|^p ds+C_1^p\varepsilon^p t^{p+1}\right)\\
&\leqslant (C_2C_3)^p t^{p-1}\int_0^t |u^\varepsilon_s-u_s|^p ds+(C_1C_2C_3)^p t^{2p}\varepsilon^p.
\end{aligned}
$$

The term in the second line has the following representation with respect to the random Poisson measure associated to Z

$$
\int_0^t [\mathfrak{F}(u^\varepsilon_{s-},v^\varepsilon_{s-})-\mathfrak{F}(u_{s-},v_{s-})]dZ_s=\int_0^t\int_{\mathbb{R}^r\setminus\{0\}}[\mathfrak{F}(u^\varepsilon_{s-},v^\varepsilon_{s-})-\mathfrak{F}(u_{s-},v_{s-})]z\tilde{N}(ds,dz).
$$

By Kunita's first inequality for the supremum of integrals with respect to the compensated random Poisson measure integrals, as stated for instance in Theorem 4.4.23 in [1], and inequality (8) we obtain

$$
\begin{aligned}
&\mathbb{E}\left[\sup_{t\in[0,T]}\Big|\int_0^t\int_{\mathbb{R}^r\setminus\{0\}}[\mathfrak{F}(u^\varepsilon_{s-},v^\varepsilon_{s-})-\mathfrak{F}(u_{s-},v_{s-})]z\tilde{N}(ds,dz)\Big|^p\right]\\
&\leqslant C_4\,\mathbb{E}\left[\left(\int_{\mathbb{R}^r\setminus\{0\}}|z|^2\nu(dz)\right)^{p/2}\left(\int_0^T|\mathfrak{F}(u^\varepsilon_{s-},v^\varepsilon_{s-})-\mathfrak{F}(u_{s-},v_{s-})|^2ds\right)^{p/2}\right]
\end{aligned}
$$

$$
\begin{aligned}
&+ C_4\, \mathbb{E}\left[\int_{\mathbb{R}^r\setminus\{0\}} |z|^p \nu(dz) \left(\int_0^T |\mathfrak{F}(u^\varepsilon_{s-}, v^\varepsilon_{s-}) - \mathfrak{F}(u_{s-}, v_{s-})|^p ds\right)\right] \\
&\leqslant C_4(C_6 C_3)^p (C_5^{p/2} T^{p/2-1} + R^{p-2} C_5) \int_0^T \mathbb{E}\left[\sup_{s\in[0,t]} |u^\varepsilon_s - u_s|^p\right] ds \\
&\quad + C_4(C_6 C_3 C_0)^p (C_5^{p/2} T^{3p/2} + R^{p-2} C_5 T^{p+1})\ \varepsilon^p \qquad (11)
\end{aligned}
$$

Since the vector fields $\mathfrak{F}$ and $(D\mathfrak{F})\mathfrak{F}$ are Lipschitz continuous, we can mimic the estimates in [7], proof of Lemma 3.1, which yields a constant $C_7 = C_7(p) > 0$, exploit that $|\Delta_s Z| \leqslant R$ almost surely, apply once again (8) and go over to the random measure representation of the quadratic variation of Z

$$
\begin{aligned}
&\Big| \sum_{0<s\leqslant t} \Phi^{\mathfrak{F}\Delta_s Z}(u^\varepsilon_{s-}, v^\varepsilon_{s-}) - \Phi^{\mathfrak{F}\Delta_s Z}(u_{s-}, v_{s-}) - (u^{\varepsilon,i}_{s-} - u^i_{s-}) - (\mathfrak{F}(u^\varepsilon_{s-}, v^\varepsilon_{s-}) - \mathfrak{F}(u_{s-}, v_{s-}))\Delta_s Z\Big|^p \\
&\quad \leqslant \Big(C_7 \sum_{0<s\leqslant t} |(u^\varepsilon_{s-} - u_{s-}, v^\varepsilon_{s-} - v_{s-})| e^{C_7|\Delta_s Z|} |\Delta_s Z|^2\Big)^p \\
&\quad \leqslant C_3^p (C_7 e^{C_47R})^p \Big(\sum_{0<s\leqslant t} (|u^\varepsilon_{s-} - u_{s-}| + |v^\varepsilon_{s-} - v_{s-}|) |\Delta_s Z|^2\Big)^p \\
&\quad = (C_3 C_7 e^{C_47R})^p \Big[\Big(\sum_{0<s\leqslant t} |u^\varepsilon_{s-} - u_{s-}| |\Delta_s Z|^2 + C_1 \varepsilon t \sum_{0<s\leqslant t} |\Delta_s Z|^2\Big)^p\Big] \\
&\quad = (C_3 C_7 e^{C_7 R})^p 2^{p-1} \Big[\Big(\int_0^t \int_{\mathbb{R}^r\setminus\{0\}} |u^\varepsilon_{s-} - u_{s-}| |z|^2 \tilde{N}(ds, dz)\Big)^p \\
&\qquad + (C_1^p \varepsilon^p t^p) \Big(\int_0^t \int_{\mathbb{R}^r\setminus\{0\}} |z|^2 \tilde{N}(ds, dz)\Big)^p\Big], \qquad (12)
\end{aligned}
$$

where $C_6 = \int_{\mathbb{R}^r\setminus\{0\}} |y|^2 \nu(dy) < \infty$. Taking the supremum and expectation for the first term of the right-hand side of (12) we apply once again Kunita's first inequality for the supremum of compensated random Poisson measures and obtain for the first summand in (12) a constant $C_8 = C_8(p) > 0$ such that

$$
\begin{aligned}
&\mathbb{E}\Big[\sup_{t\in[0,T]} \Big(\int_0^t \int_{\mathbb{R}^r\setminus\{0\}} |u^\varepsilon_{s-} - u_{s-}| |z|^2 \tilde{N}(ds, dz)\Big)^p\Big] \\
&\leqslant R^p \mathbb{E}\Big[\sup_{t\in[0,T]} \Big(\int_0^t \int_{\mathbb{R}^r\setminus\{0\}} |u^\varepsilon_{s-} - u_{s-}| |z| \tilde{N}(ds, dz)\Big)^p\Big] \\
&\leqslant R^p C_8 \Big(\mathbb{E}\Big[\Big(\int_0^T \int_{\mathbb{R}^r\setminus\{0\}} |u^\varepsilon_{s-} - u_{s-}|^2 |y|^2 \nu(dy) ds\Big)^{\frac{p}{2}}\Big] + \mathbb{E}\Big[\int_0^T \int_{\mathbb{R}^r\setminus\{0\}} |u^\varepsilon_{s-} - u_{s-}|^p |y|^p \nu(dy) ds\Big]\Big) \\
&= R^p C_8 C_5^{p/2} T^{p/2-1} \mathbb{E}\left[\sup_{t\in[0,T]} \int_0^t |u^\varepsilon_{s-} - u_{s-}|^p ds\right] + C_8 R^{2p-2} C_5 \mathbb{E}\left[\left(\int_0^T \sup_{s\in[0,t]} |u^\varepsilon_{s-} - u_{s-}|^p dt\right)\right] \\
&= \Big(R^p C_8 C_5^{p/2} T^{p/2-1} + C_8 R^{2p-2} C_5\Big) \int_0^T \mathbb{E}\left[\sup_{s\in[0,t]} |u^\varepsilon_{s-} - u_{s-}|^p\right] dt. \qquad (13)
\end{aligned}
$$

For the last term on the right-hand side of (12) there exists a constant $C_9(p) > 0$ satisfying

$$
\begin{aligned}
&\mathbb{E}\left[\sup_{t\in[0,T]}\Big(\int_0^t\int_{\mathbb{R}^r\setminus\{0\}}|z|^2\tilde N(ds,dz)\Big)^p\right]\\
&\qquad\leqslant C_9T^p\Big(\int_{\mathbb{R}^r\setminus\{0\}}|z|^2\nu(dz)\Big)^p+C_9T\int_{\mathbb{R}^r\setminus\{0\}}|z|^p\nu(dz)=C_9T^pC_5^p+C_9R^{p-2}TC_5. \quad (14)
\end{aligned}
$$

Taking the supremum and expectation in inequality (10) we combine (11), (13) and (14) and obtain constants C_{10} and C_{11}

$$
\begin{aligned}
&\mathbb{E}\left[\sup_{t\in[0,T]}|u_t^\varepsilon-u_t|^p\right]\\
&\leqslant C_{10}(T^{2p}+T^{p+1})\varepsilon^p+C_{11}\Big(T^{p-1}+1\Big)\int_0^T\mathbb{E}\left[\sup_{s\in[0,t]}|u_s^\varepsilon-u_s|^p\right]dt\\
&=: a_\varepsilon(T)+b(T)\int_0^T\mathbb{E}\left[\sup_{s\in[0,t]}|u_s^\varepsilon-u_s|^p\right]dt.
\end{aligned}
$$

A standard integral version of Gronwall's inequality, as stated for instance in Lemma D.2 in [10], yields that

$$
\begin{aligned}
\mathbb{E}\left[\sup_{t\in[0,T]}|u_t^\varepsilon-u_t|^p\right]&\leqslant a_\varepsilon(T)(1+b(T)T\exp(b(T)T))\\
&\leqslant C_{10}T^{p+1}(1+T^{p-1})\varepsilon^p\Big[1+C_{11}T(1+T^{p-1})\exp\big(C_{11}T(1+T^{p-1})\big)\Big]\\
&\leqslant C_{12}T^p(1+T^p)^2\varepsilon^p\exp\left(C_{13}T^p\right).
\end{aligned}
$$

Hence

$$
\begin{aligned}
\left(\mathbb{E}\left[\sup_{t\in[0,T]}|u_t^\varepsilon-u_t|^p\right]\right)^{\frac1p}&\leqslant C_{12}\,\varepsilon\,T(1+T)^2\exp\left(C_{13}\,T^p\right)\\
&\leqslant C_{12}\,\varepsilon\,T\exp\left(C_{14}\,T^p\right). \qquad (15)
\end{aligned}
$$

Eventually Minkowski's inequality and the estimates (5), (8) and (15) yield the desired result

$$
\begin{aligned}
&\left(\mathbb{E}\left[\sup_{s\leqslant T\wedge\tau^\varepsilon}|\Psi(X_t^\varepsilon(x_0))-\Psi(X_t(x_0))|^p\right]\right)^{\frac1p}\\
&\quad\leqslant C'\left(\mathbb{E}\left[\sup_{s\leqslant T\wedge\tau^\varepsilon}|u_s^\varepsilon-u_s|^p\right]\right)^{\frac1p}+C'\left(\mathbb{E}\left[\sup_{s\leqslant T\wedge\tau^\varepsilon}|v_s^\varepsilon-v_s|^p\right]\right)^{\frac1p}\\
&\quad\leqslant C'C_{14}\,\varepsilon\,T\exp\left(C_{13}\,T^p\right)+C'C_1T\varepsilon\\
&\quad\leqslant C_{15}\,\varepsilon\,T\exp\left(C_{13}\,T^p\right).
\end{aligned}
$$

This finishes the proof. ■

If Z has a continuous component, the solution of equation (1) also contains a continuous Stratonovich component, see [7]. Combining the proof above with the proof of Lemma 2.1 in [4] we conclude the following.

Corollary 2.1 *Let Z be a Lévy process with characteristic triplet (b, ν, A) in $\mathbb{R}^r$ for a drift vector $b \in \mathbb{R}^r$ and the covariance matrix A and ν as in Lemma 2.1. Then estimate (4) of Proposition 2.1 holds true with an appropriate choice of the constants K_1 and K_2.*

Proof: The estimates of $\mathbb{E}\left[\sup_{t\in[0,T]} |u_t^\varepsilon - u_t|^p\right]$ in both cases -continuous and pure jumps- just before applying Gronwall's lemma yields polynomial estimates in T and ε of the same degree. Hence Gronwall's inequality guarantees the same estimates modulo constants. ■

3 Averaging functions on the leaves

We assume that each leaf $L_q \in \mathfrak{M}$ passing through $q \in M$ contains a unique ergodic invariant measure μ_q of the unperturbed foliated system (1) with initial condition $x_0 = q$.

Let $\Psi : M \to \mathbb{R}$ be a differentiable function. We define the average of Ψ with respect to μ_q in the following way. If v is the vertical coordinate of q, that is $q = \varphi(v, u)$, we define

$$Q^\Psi(v) := \int_{L_q} \Psi(y)\mu_q(dy). \tag{16}$$

and fix the notation $\Pi(q) := \pi_2(\varphi^{-1}(q)) = v$, where π_2 is the projection on $V \subset \mathbb{R}^d$. We assume that $Q^\Psi(\Pi(X_s^\varepsilon))$ is Riemann inegrabile of with respect to s.

For fixed $x_0 \in M$ and $\varepsilon > 0$ we denote by $\tau = \tau_{x_0}^\varepsilon$ the stopping time, which stops when $X_\cdot^\varepsilon(x_0)$ leaves the open neighbourhood $U \subset M$, which is diffeomorphic to $L_{x_0} \times V$.

Proposition 3.1 *Given $\Psi : M \to \mathbb{R}$ differentiable and $Q^\Psi : V \to \mathbb{R}$ its average on the leaves given by (16). For $t \geqslant 0$ we denote by*

$$\delta^\Psi(\varepsilon, t) := \int_0^{t\wedge\varepsilon\tau} \Psi(X_{\frac{r}{\varepsilon}}^\varepsilon(x_0)) - Q^\Psi(\Pi(X_{\frac{r}{\varepsilon}}^\varepsilon(x_0)))dr.$$

Then $\delta^\Psi(\varepsilon, t)$ tends to zero, when t or ε tend to zero. Moreover, if Q^Ψ is α-Hölder continuous with $\alpha > 0$, then for $p \geqslant 1$ and any $\beta \in (0, \frac{1}{2})$ holds the estimate

$$\left(\mathbb{E}\left[\sup_{s\leqslant t} |\delta^\Psi(\varepsilon, s)|^p\right]\right)^{\frac{1}{p}} \leqslant \sqrt{t}|\ln(\varepsilon)|^{-\frac{\beta}{p}} h(t, \varepsilon),$$

where $h(t, \varepsilon)$ is continuous in (t, ε) and tends to zero when t or ε do so.

Proof: (First part.) For ε sufficiently small and $t \geqslant 0$ we define the partition

$$t_0 = 0 < t_1^\varepsilon < \cdots < t_{N^\varepsilon}^\varepsilon \leqslant \frac{t}{\varepsilon} \wedge \tau$$

as long as X^ε has not left U with the following step size

$$h_\varepsilon := t|\ln\varepsilon|^{2\frac{\beta}{p}}$$

by

$$t_n^\varepsilon := nh_\varepsilon \qquad \text{for} \qquad 0 \leqslant n \leqslant N^\varepsilon \qquad \text{where} \qquad N^\varepsilon = \lfloor(\varepsilon|\ln\varepsilon|^{2\frac{\beta}{p}})^{-1}\rfloor.$$

We now represent the first summand of δ^Ψ by

$$\int_0^{t\wedge\varepsilon\tau} \Psi(X_{\frac{r}{\varepsilon}}^\varepsilon(x_0))dr = \varepsilon\int_0^{\frac{t}{\varepsilon}\wedge\tau} \Psi(X_r^\varepsilon(x_0))dr$$

$$= \varepsilon \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} \Psi(X_r^\varepsilon(x_0)) + \varepsilon \int_{t_n}^{\frac{t}{\varepsilon}\wedge\tau} \Psi(X_r^\varepsilon(x_0))dr$$

We lighten notation and omit for convenience in the sequel all super and subscript ε and Ψ as well as the initial value x_0. We denote by θ the canonical shift operator on the canonical probability space $\Omega = D(\mathbb{R}, M)$ of càdlàg functions. Let $F_t(\cdot, \omega)$ the stochastic flow of the original unpertubed system in M. The triangle inequality yields

$$|\delta^\Psi(\varepsilon, t)| \leqslant |A_1(t, \varepsilon)| + |A_2(t, \varepsilon)| + |A_3(t, \varepsilon)| + |A_4(t, \varepsilon)|, \tag{17}$$

where

$$A_1(t, \varepsilon) := \varepsilon \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} [\Psi(X_r^\varepsilon) - \Psi(F_{r-t_n}(X_{t_n}^\varepsilon, \theta_{t_n}(\omega)))] \, dr$$

$$A_2(t, \varepsilon) := \varepsilon \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} [\Psi(F_{r-t_n}(X_{t_n}^\varepsilon, \theta_{t_n}(\omega))) - hQ(\Pi(X_{t_n}^\varepsilon))] \, dr$$

$$A_3(t, \varepsilon) := \sum_{n=0}^{N-1} \varepsilon h Q(\Pi(X_{t_n}^\varepsilon)) - \int_0^{t\wedge\varepsilon\tau} Q(\Pi(X_{\frac{r}{\varepsilon}}^\varepsilon)) \, dr$$

$$A_4(t, \varepsilon) := \varepsilon \int_{t_n}^{\frac{t}{\varepsilon}\wedge\tau} \Psi(X_r^\varepsilon(x_0))dr.$$

The following four lemmas estimate the preceding terms. This being done the proof is finished.

Lemma 3.1 *For any $\gamma \in (0, 1)$ there exists a function $h_1 = h_1(\gamma)$ such that*

$$\left(\mathbb{E}\left[\sup_{s\leqslant t} |A_1(s, \varepsilon)|^p\right]\right)^{\frac{1}{p}} \leqslant K_1 t \varepsilon^\gamma h_1(t, \varepsilon),$$

where h_1 is continuous in ε and t and tends to zero when ε and t do so.

Proof: The proof is identical to Lemma 3.2 in [4], since Proposition 2.1 provides the same asymptotic bounds as Lemma 2.1 in [4], which enters here. Furthermore, only the Markov property of the solutions of equation (3) is exploited. ■

Lemma 3.2 *For the process A_2 in inequality (17) there exists a constant $K_3 > 0$ such that*

$$\left(\mathbb{E}\left[\sup_{s\leqslant t} |A_2(s, \varepsilon)|^p\right]\right)^{\frac{1}{p}} \leqslant K_3 \sqrt{t} |\ln(\varepsilon)|^{-\frac{\beta}{p}}.$$

Proof: The proof is identical to the proof of Lemma 3.3 in [4], since it only exploits the Markov property, the existence of first moments and the rate of convergence of the strong law of large numbers. ■

Lemma 3.3 *Assume that Q^Ψ is α-Hölder continuous with $\alpha > 0$. Then the process A_3 in inequality (17) satisfies*

$$\left(\mathbb{E}\left[\sup_{s\leqslant t} |A_3(s, \varepsilon)|^p\right]\right)^{\frac{1}{p}} \leqslant K_4 t^{1+\alpha} \varepsilon^\alpha |\ln(\varepsilon)|^{\frac{2\alpha\beta}{p}}$$

for a positive constant $K_4 > 0$.

Proof: We lighten notation $Q = Q^\Psi$. We consider the interval $[0,t]$ with the partition $0 < \varepsilon t_1 < \cdots < \varepsilon t_N \leqslant t$

$$\begin{aligned}
|A_3(t,\varepsilon)| &= \Big| \sum_{n=0}^{N-1} \varepsilon h Q(\Pi(X^\varepsilon_{t_n})) - \int_0^{t\wedge\varepsilon\tau} Q(\Pi(X^\varepsilon_{\frac{r}{\varepsilon}}))\, dr \Big| \\
&\leqslant \varepsilon \sum_{n=0}^{N-1} h \sup_{\varepsilon t_n \leqslant s < \varepsilon t_{n+1}} |Q(\Pi(X^\varepsilon_s)) - Q(\Pi(X^\varepsilon_{t_n}))| \\
&\leqslant \varepsilon C_1 h N \sup_{\varepsilon t_n \leqslant s < \varepsilon t_{n+1}} |v^\varepsilon_s - v^\varepsilon_{t_n}|^\alpha \\
&\leqslant \varepsilon C_2 h N (\varepsilon h)^\alpha \\
&\leqslant C_3 \varepsilon^{1+\alpha} t^{1+\alpha} |\ln(\varepsilon)|^{(1+\alpha)\frac{2\beta}{p}} \varepsilon^{-1} |\ln(\varepsilon)|^{-\frac{2\beta}{p}} \\
&= C_3 t^{1+\alpha} \varepsilon^\alpha |\ln(\varepsilon)|^{\frac{2\alpha\beta}{p}}.
\end{aligned}$$

■

Lemma 3.4 *The process A_4 satisfies*

$$\mathbb{E}\Big[\sup_{s\leqslant t} |A_4(s,\varepsilon)|^p\Big] \leqslant K_5 t\varepsilon |\ln\varepsilon|^{\frac{2\beta}{p}},$$

where $K_5 = \|\Psi\|_{\infty,U}$.

The proof is identical to the Proof of Lemma 3.5 in [4].

(Final step of Proposition 3.1) Collecting the results of the previous lemmas yields with the help of Minkowski's inequality the desired result

$$\begin{aligned}
\left(\mathbb{E}\left[\sup_{s\leqslant t} |\delta^\Psi(\varepsilon,s)|^p\right]\right)^{\frac{1}{p}} &\leqslant t^{\frac{1}{2}} |\ln\varepsilon|^{-\frac{\beta}{p}} \left(K_1 t^{\frac{1}{2}} \varepsilon^\gamma h_1(\varepsilon,t) + K_3 + K_4 t^{\frac{1}{2}+\alpha} \varepsilon^\alpha |\ln(\varepsilon)|^{\frac{(2\alpha-1)\beta}{p}} + K_5 t^{\frac{1}{2}} |\ln\varepsilon|^{\frac{3\beta}{p}} \right) \\
&=: t^{\frac{1}{2}} |\ln\varepsilon|^{-\frac{\beta}{p}} h(t,\varepsilon),
\end{aligned}$$

where $h(t,\varepsilon)$ tends to zero if ε or t does so. ■

4 An averaging principle

For the main result we are going to assume the following
Hypothesis H: For any Lipschitz continuous function Ψ on M, its corresponding average function Q^Ψ on the transversal space V, which indexes the leaves is also Lipschitz.

We use the derivatives of each component of $\Pi : M \to V \subset \mathbb{R}^d$, where

$$\Pi(\cdot) = (\Pi_1(\cdot), \dots, \Pi_d(\cdot)),$$

in order to get the averages $Q^{d\pi_i(K)}(x)$ of the real functions $d\Pi_i(K)$, $i = 1, \dots, d$ on each leaf of L_x.

Theorem 4.1 *Assume that the unperturbed foliated system (1) on M satisfies Hypothesis H. Let w be the solution of the deterministic ODE in the transversal component $V \subset \mathbb{R}^n$.*

$$\frac{dw}{dt}(t) = (Q^{d\Pi_1(K)}, \ldots, Q^{d\Pi_d(K)})\,(w(t))\,, \tag{18}$$

with initial condition $w(0) = \Pi(x_0) = 0$. Let T_0 be the time that $w(t)$ reaches the boundary of V. Then we have that:

1. *For all $0 < t < T_0$, $\beta \in (0, \frac{1}{2})$ and $2 \leqslant p < \infty$*

$$\left(\mathbb{E}\left[\sup_{s\leqslant t} \|\Pi\big(X^{\varepsilon}_{\frac{s\wedge\tau}{\varepsilon}}\big) - w(s)\|^p\right]\right)^{\frac{1}{p}} \leqslant \sqrt{t}|\ln(\varepsilon)|^{-\frac{\beta}{p}} h(t,\varepsilon),$$

where $h(t,\varepsilon)$ is continuous and converges to zero, when ε or t do so.

2. *For $\gamma > 0$, let*

$$T_\gamma := \inf\{t > 0 \mid \operatorname{dist}(w(t), \partial V) \leqslant \gamma\}.$$

The exit times of the two systems satisfy the estimates

$$\mathbb{P}(\varepsilon\tau < T_\gamma) \leqslant \gamma^{-p} t^{\frac{p}{2}} |\ln\varepsilon|^{-\beta} h(t,\varepsilon)^p.$$

The second part of the theorem above guarantees the robustness of the averaging phenomenon in transversal direction.

Proof: The gradient of each Π_i is orthogonal to the leaves. Hence by Itô's formula for canconical Marcus integrals, see e.g. [7] Proposition 4.2, we obtain for $i = 1, \ldots, d$ that

$$\Pi_i\big(X^{\varepsilon}_{\frac{t\wedge\tau}{\varepsilon}}\big) = \int_0^{\frac{t\wedge\tau}{\varepsilon}} d\Pi_i(\varepsilon K)(X^{\varepsilon}_r)dr = \int_0^{t\wedge\tau} d\Pi_i(K)(X^{\varepsilon}_{\frac{r}{\varepsilon}})dr. \tag{19}$$

We may continue and change the variable

$$\begin{aligned}
|\Pi_i\big(X^{\varepsilon}_{\frac{t\wedge\tau}{\varepsilon}}\big) - w_i\big(t\big)| &\leqslant \int_0^{t\wedge\tau} |Q^{d\Pi_i(K)}(X^{\varepsilon}_{\frac{r}{\varepsilon}}) - Q^{d\Pi_i(K)}(w(r))|dr + |\delta^{d\Pi_i}(t,\varepsilon)| \\
&\leqslant C_1 \int_0^{t\wedge\tau} |\Pi_i(X^{\varepsilon}_{\frac{r}{\varepsilon}}) - w_i(r)|dr + |\delta^{d\Pi_i}(t,\varepsilon)| \\
&\leqslant C_2 \int_0^{t\wedge\tau} |\Pi(X^{\varepsilon}_{\frac{r}{\varepsilon}}) - w(r)|dr + \sum_{i=1}^{N} |\delta^{d\Pi_i}(t,\varepsilon)|
\end{aligned}$$

Since the right-hand side does not depend on i, we can sum over i at the left-hand side and apply Gronwall's lemma. This yields

$$|\Pi\big(X^{\varepsilon}_{\frac{t\wedge\tau}{\varepsilon}}\big) - w\big(t\big)| \leqslant e^{C_2 t} \sum_{i=1}^{d} |\delta^{d\Pi_i}(t,\varepsilon)|.$$

An application of Proposition 3.1 finishes the proof of the first statement. For the second part we calculate with the help of Chebyshev's inequality

$$\mathbb{P}\big(\varepsilon\tau < T_\gamma\big) = \mathbb{P}\big(\sup_{s\leqslant T_\gamma\wedge\varepsilon\tau} |\Pi(X^{\varepsilon}_{\frac{t\wedge\tau}{\varepsilon}}) - w(s)| > \gamma\big)$$

$$\leqslant \gamma^{-p}\mathbb{E}\left[\sup_{s\leqslant T_\gamma\wedge\varepsilon\tau}|\Pi(X^{\varepsilon}_{\frac{t\wedge\tau}{\varepsilon}})-w(s)|^p\right]$$
$$\leqslant \gamma^{-p}t^{\frac{p}{2}}|\ln(\varepsilon)|^{-\beta}h^p(t,\varepsilon).$$

■

Example. As a simple but illustrative example of the phenomenon we consider a foliation (cf. [4]) whose transversal space is richer than the leaves themselves, hence the range of the impact for different perturbations is relatively large. We consider $M=\mathbb{R}^3\setminus\{(0,0,z),z\in\mathbb{R}\}$ with the 1-dimension horizontal circular foliation of M where the leaf passing through a point $q=(x,y,z)$ is given by the horizontal circle

$$L_p=\{(\sqrt{x^2+y^2}\cos\theta,\sqrt{x^2+y^2}\sin\theta,z),\quad \theta\in[0,2\pi)\}.$$

For an initial condition $q_0=(x_0,y_0,z_0)$, say with $x_0\geqslant 0$ consider the local foliated coordinates in the neighbourhood $U=\mathbb{R}^3\setminus\{(x,0,z);x\leqslant 0;z\in\mathbb{R}\}$ given by cylindrical coordinates. Hence, using the same notation as before $\varphi=(u,v)$ will be defined by $\varphi: U\subset M\to(-\pi,\pi)\times\mathbb{R}_{>0}\times\mathbb{R}$, where $u\in(-\pi,\pi)$ is angular and $v=(r,z)\in\mathbb{R}_{>0}\times\mathbb{R}$ such that $\varphi^{-1}:(u,v)\mapsto(r\cos u,r\sin u,z)\in M$. In this coordinates system, the transversal projections Π_1 and Π_2 correspond to the radial r-component and the vertical z-coordinate, respectively. Consider the foliated linear SDE on M consisting of random rotations:

$$dX_t=\Lambda X_t\ \diamond dZ_t,\qquad X_0=q_0=(x_0,y_0,z_0), \tag{20}$$

where

$$\Lambda=\begin{pmatrix}0&-1&0\\1&0&0\\0&0&0\end{pmatrix}.$$

For convenience the process $Z=(Z_t)_{t\geqslant 0}$ is a symmetric Lévy process in $\mathbb{R}$ with characteristic triplet $(a,\nu,1)$, where ν is a Lévy jump measure with support $[-1,1]$ and satisfies the integrability condition $\int_{\mathbb{R}}|z|^2\nu(dz)<\infty$, that is, by its Lévy-Itô decomposition it has almost surely the shape

$$Z_t=at+B_t+\int_0^t\int_{[-1,1]}y\tilde{N}(dsdy),$$

where B is a standard Brownian motion independent of the pure jump part and $\tilde{N}$ is the compensated random Poisson measure with intensity measure $dt\otimes\nu$. Equation (20) is defined as follows: First note that $\Lambda^2=\operatorname{diag}(-1,-1,0)$. Secondly, note that for $z\in\mathbb{R}$, $z\neq 0$ the solution flow Φ of the equation

$$\frac{d}{d\sigma}Y(\sigma)=F(Y(\sigma))z,\qquad Y(0)=q\qquad \text{where } F(\bar{q})=\Lambda\bar{q}$$

is obtained by a simple calculation as

$$\Phi^{Fz}(q)=Y(1;q)=\begin{pmatrix}x\cos(z)-y\sin(z)\\x\sin(z)+y\cos(z)\\z\end{pmatrix}.$$

Note that here the compromise between the jump increments, the vector field and the geometry is satisfied in the sense of inclusion (4), since the norm of the vector fields $\Phi^{F(x)}(x)-F(x)z$ decreases accordingly when the radius of injectivity diminishes. Hence we obtain

$$X_t=q_0+\int_0^t(a\Lambda X_s+\frac{1}{2}\Lambda^2X_s)ds+\int_0^t\Lambda X_{s-}(dB_s+z\tilde{N}(ds,dz))$$

$$+\sum_{0<s\leqslant t}(\Phi^{F\Delta_s Z}(X_{s-})-X_{s-}-F(X_{s-})\Delta_s Z).$$

We ignore from now on the constant third component $X_t^3=z_0$ for all $t\geqslant 0$ almost surely and keep the name X for (X^1,X^2) for convenience. Using the chain rule of the Marcus integral, as stated in Proposition 4.2 in [7], we verify for $\eta(x,y):=x^2+y^2$ that for $X=(X^1,X^2)$

$$d\eta(X_t^1,X_t^2)=-2X_{t-}\Lambda X_{t-}\diamond dZ_t=0.$$

The invariant measures μ_q in the leaves L_q passing through points $q\in M$ are given by normalized Lebesgue measures in the circle L_q, hence Hypothesis (H) is trivially satisfied. We investigate the effective behaviour of a small transversal perturbation of order ε:

$$dX_t^\varepsilon=\Lambda X_t^\varepsilon\ \diamond dZ_t+\varepsilon K(X_t^\varepsilon)\ dt.$$

with initial condition $q_0=(1,0,0)$. In this example we shall consider two classes of perturbing vector field K.

(A) Constant perturbation εK. Assume that the perturbation is given by a vector field which is constant $K=(\kappa_1,\kappa_2,\kappa_3)$ with respect to Euclidean coordinates in M. Then, the average horizontal component $Q^{d\Pi_1 K}=0$ and the vertical z-component is constant $Q^{d\Pi_2 K}=\kappa_3$. Hence the transversal component in Theorem 4.1 for initial condition $q_0=(1,0,0)$ is given by $w(t)=(1,\kappa_3 t)$ for all $t\geqslant 0$. Theorem 4.1 establishes a minimum rate of convergence to zero of the difference between each of the transversal components. Hence for the radial component of the perturbed systems $w_1(t)\equiv 1$ and $\Pi_1(X^\varepsilon_{\frac{t\wedge\tau^\varepsilon}{\varepsilon}})$ holds that, for $p\geqslant 2$

$$\left[\mathbb{E}\left(\sup_{s\leqslant t}\left|\Pi_1(X^\varepsilon_{\frac{t\wedge\tau^\varepsilon}{\varepsilon}})-1\right|^p\right)\right]^{\frac{1}{p}}$$

goes to zero as ε or t goes to zero with the prescribed rate of convergence. We have that

$$X^\varepsilon_{\frac{t}{\varepsilon}}=\begin{pmatrix}\cos\left(\frac{at}{\varepsilon}+B_{\frac{t}{\varepsilon}}+\int_0^{\frac{t}{\varepsilon}}\int_{[-1,1]}y\tilde N(dsdy)\right)\\ \sin\left(\frac{at}{\varepsilon}+B_{\frac{t}{\varepsilon}}+\int_0^{\frac{t}{\varepsilon}}\int_{[-1,1]}y\tilde N(dsdy)\right)\\ 0\end{pmatrix}$$

$$+\varepsilon\begin{pmatrix}\kappa_1\sin\left(\frac{at}{\varepsilon}+B_{\frac{t}{\varepsilon}}+\int_0^{\frac{t}{\varepsilon}}\int_{[-1,1]}y\tilde N(dsdy)\right)+\kappa_2\cos\left(\frac{at}{\varepsilon}+B_{\frac{t}{\varepsilon}}+\int_0^{\frac{t}{\varepsilon}}\int_{[-1,1]}y\tilde N(dsdy)\right)-\kappa_2\\ -\kappa_1\cos\left(\frac{at}{\varepsilon}+B_{\frac{t}{\varepsilon}}+\int_0^{\frac{t}{\varepsilon}}\int_{[-1,1]}y\tilde N(dsdy)\right)+\kappa_2\sin\left(\frac{at}{\varepsilon}+B_{\frac{t}{\varepsilon}}+\int_0^{\frac{t}{\varepsilon}}\int_{[-1,1]}y\tilde N(dsdy)\right)-\kappa_1\\ 0\end{pmatrix}$$

By normalization and using the symmetry, one can fix any k_1 and k_2; for simplicity, we shall fix $K=(1,0,0)$, hence, in this case, for $t\leq\tau^\varepsilon$

$$r(t)=1+\varepsilon^2\left[2+\cos\left(\frac{at}{\varepsilon}+B_{\frac{t}{\varepsilon}}+\int_0^{\frac{t}{\varepsilon}}\int_{[-1,1]}y\tilde N(dsdy)\right)\right]$$
$$-\varepsilon\left[\cos 2\left(\frac{at}{\varepsilon}+B_{\frac{t}{\varepsilon}}+\int_0^{\frac{t}{\varepsilon}}\int_{[-1,1]}y\tilde N(dsdy)\right)+\sin\left(\frac{at}{\varepsilon}+B_{\frac{t}{\varepsilon}}+\int_0^{\frac{t}{\varepsilon}}\int_{[-1,1]}y\tilde N(dsdy)\right)\right].$$

Hence, the comparison of the second transversal component

$$|\Pi_2\left(X^{\varepsilon}_{\frac{t\wedge\tau^{\varepsilon}}{\varepsilon}}\right) - w_2(t)| \equiv 0$$

for all $t \geqslant 0$ and the convergence of the theorem is trivially verified.

(B) Linear perturbation $\varepsilon K(x,y,z) = \varepsilon(x,0,0)$. For the sake of simplicity, consider a one dimensional and horizontal linear perturbation, which in this case can be written in the form $K(x,y,z) = (x,0,0)$. The z-coordinate average vanishes trivially. For the radial component, we have that $d\Pi_1 K(q) = r\cos^2(u)$, where u is the angular coordinate of q whose distance to the z-axis (radial coordinate) is r. Hence the average with respect to the invariant measure on the leaves is given by $Q^{d\Pi_1 K} = r/2$ for leaves with radius r.

For initial value $q_0 = (x_0,y_0,z_0) = (r_0\cos(u_0), r_0\sin(\theta_0), z_0)$ the transversal system stated in Theorem 4.1 is then $w(t) = (e^{\frac{t}{2}}r_0, z_0)$. Hence the result guarantees that the radial part $\Pi_1\big(X^{\varepsilon}_{\frac{t\wedge\tau^{\varepsilon}}{\varepsilon}}\big)$ must have a behaviour close to the exponential $e^{\frac{t}{2}}$ in the sense that

$$\left[\mathbb{E}\left(\sup_{s\leqslant t}\left|\Pi_1\big(X^{\varepsilon}_{\frac{t\wedge\tau^{\varepsilon}}{\varepsilon}}\big) - e^{\frac{t}{2}}\right|^p\right)\right]^{\frac{1}{p}}$$

goes to zero when ε goes to zero. The fundamental solution of linearly perturbed Marcus equation is given by the exponential of the matrix

$$\begin{pmatrix} \varepsilon t & -(at + B_t + \int_0^t \int_{[-1,1]} y\tilde{N}(dsdy)) & 0 \\ (at + B_t + \int_0^t \int_{[-1,1]} y\tilde{N}(dsdy)) & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

In fact, the eigenvalues for the first two coordinates (the horizontal plane) are

$$\lambda_{1,2} := \left(\frac{\varepsilon t}{2} \pm \frac{1}{2}\sqrt{\varepsilon^2 t^2 - 4(at + B_t + \int_0^t \int_{[-1,1]} y\tilde{N}(dsdy))^2}\right),$$

whose real part is given by $\varepsilon/2$, with probability increasing to 1 as ε goes to 0. This result points out that the top Lyapunov exponent (in the horizontal directions) of the perturbed system in the original time scale is given by $\varepsilon/2$.

Acknowledgements: The first author would like to express his gratitude to Berlin Mathematical School (BMS), International Research Training Group (IRTG) 1740: Dynamical Phenomena in Complex Networks: Fundamentals and Applications and the probability group at Humboldt-Universität zu Berlin and Universität Potsdam for various infrastructural support. This article has been written while the second author is in a sabbatical visit to Humboldt-Universität zu Berlin. He would like to express his gratitude to Prof. Peter Imkeller and his research group for the nice and friendly hospitality. Partially supported by FAPESP 11/50151-0 and 12/03992-1.

References

[1] D. Applebaum. Levy processes and stochastic calculus. *Cambridge university press*, 2nd edition, 2009.

[2] V. Arnold – *Mathematical Methods in Classical Mechanics*. 2nd ed. Berlin, Springer (1989).

[3] A. Borodin and M. Freidlin – Fast oscillating random perturbations of dynamical systems with conservation laws. *Ann. Inst. H. Poincaré. Prob. Statist.*, 31 (1995) 485-525.

[4] I. I. G. Gargate and P. R. Ruffino – *An averaging principle for diffusions in foliated spaces*, arxiv:1212.1587v4 (math.DS), 2013.

[5] Y. Kabanov and S. Pergamenshchikov –*Two-Scale Stochastic Systems: asymptotic analysis and control*. Springer-Verlag, 2003.

[6] R. Khasminski and N. Krylov – On averaging principle for diffusion processes with null-recurrent fast component. *Stoch Proc Appl.* 93 (2001) 229-240.

[7] T. G. Kurtz and E. Pardoux and Philip Protter – *Stratonovich stochastic differential equations driven by general semimartingales* Annales de l'H.I.P., section B, 1995, 31(2), 351-377.

[8] Xue-Mei Li – An averaging principle for a completely integrable stochastic Hamiltonian systems. *Nonlinearity*, 21 (2008) 803-822.

[9] Ph. Protter – Stochastic integration and differential equations. *Springer-Verlag, Berlin Heidelberg New York*, (2004).

[10] S. G. Sell and Y. You. Dynamics of evolutionary equations. *Applied Mathematical Sciences 143*, Springer, 2002.

[11] J.A. Sanders, F. Verhulst and J. Murdock – *Averaging Methods in Nonlinear dynamical Systems*. 2nd Ed. Springer, 2007.

[12] P. Tondeur. *Foliations on Riemannian manifolds*. Universitext, Springer Verlag, Berlin-Heidelberg-New York, 1988.

[13] P. Walcak – *Dynamics of foliations, groups and pseudogroups*. Birkhäuser Verlag 2004.