Homogenisation for anisotropic kinetic random motions[*]

Pierre Perruchaud[†]

Abstract

We introduce a class of kinetic and anisotropic random motions $(x_t^\sigma, v_t^\sigma)_{t \geq 0}$ on the unit tangent bundle $T^1 \mathcal{M}$ of a general Riemannian manifold $(\mathcal{M}, g)$, where σ is a positive parameter quantifying the amount of noise affecting the dynamics. As the latter goes to infinity, we then show that the time rescaled process $(x_{\sigma^2 t}^\sigma)_{t \geq 0}$ converges in law to an explicit anisotropic Brownian motion on $\mathcal{M}$. Our approach is essentially based on the strong mixing properties of the underlying velocity process and on rough paths techniques, allowing us to reduce the general case to its Euclidean analogue. Using these methods, we are able to recover a range of classical results.

Keywords: diffusion processes; homogenization; mixing; rough paths theory; Riemannian manifolds.
AMS MSC 2010: 58J65; 60H10; 60G51.
Submitted to EJP on November 21, 2018, final version accepted on March 2, 2020.
Supersedes arXiv:1811.08415.
Supersedes HAL:hal-01930183.

1 Introduction

We consider a class of anisotropic and kinetic random motions on the unit tangent space of a general Riemannian manifold $(\mathcal{M}, g)$ of dimension $d \geq 2$. In the simplest case when the base manifold is the Euclidean space $\mathbb{R}^d$, the typical process we have in mind can be described as follows: let $\sigma > 0$ be a positive parameter and let $(B_t)_{t \geq 0}$ be a Brownian motion in $\mathbb{R}^d$ with (non identity) covariance matrix $\Sigma = A^* A$. We construct an anisotropic diffusion process $(v_t)_{t \geq 0} = (v_t^\sigma)_{t \geq 0}$ on the Euclidean sphere $\mathbb{S}^{d-1} \subset \mathbb{R}^d$ by solving the Stratonovich differential equation

$$\mathrm{d}v_t = \sigma \Pi_{v_t^\perp} \circ \mathrm{d}B_t, \tag{1.1}$$

---

[*]Submitted while the author was affiliated with the Université de Rennes 1, IRMAR - UMR 6625, France.
[†]Department of Mathematics, University of Notre Dame, Indiana, USA.
 E-mail: pperruch@nd.edu

where $\Pi_{v_t^\perp}$ denotes the projection on the orthogonal of v_t. We then integrate the velocity process $(v_t)_{t\geq 0}$ to obtain a process $(x_t)_{t\geq 0} = (x_t^\sigma)_{t\geq 0}$ with values in $\mathbb{R}^d$

$$x_t := x_0 + \int_0^t v_s \mathrm{d}s. \qquad (1.2)$$

The process $(x_t, v_t)_{t\geq 0}$ is thus a diffusion process with values in the unit tangent space $T^1\mathbb{R}^d = \mathbb{R}^d \times \mathbb{S}^{d-1}$. The first projection $(x_t)_{t\geq 0}$ is a $\mathcal{C}^1$ curve in $\mathbb{R}^d$, which inherits the anisotropy of the velocity process $(v_t)_{t\geq 0}$, and the positive parameter σ allows one to slow or speed up the clock of the latter. The next figure shows an approximation of a sample path of the resulting process.
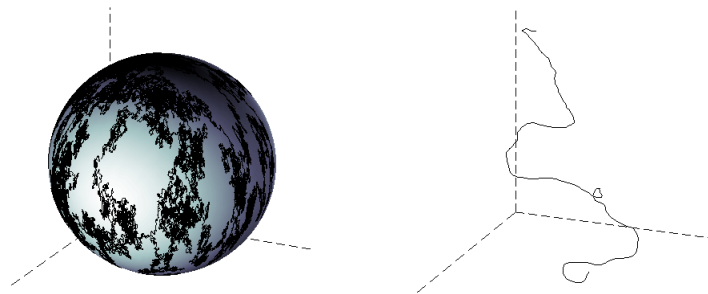

Figure 1: A sample path of the velocity process $(v_t)_{0 \leq t \leq 10}$ on $\mathbb{S}^2$ (left) and the corresponding $(x_t)_{0 \leq t \leq 10}$ in $\mathbb{R}^3$ (right) for the choice of covariance matrix $\Sigma = \mathrm{diag}(1, 1.1, 1.2)$.

On a general Riemannian manifold $(\mathcal{M}, g)$, an analogue process $(x_t, v_t)_{t\geq 0}$ with values in the unit tangent bundle $T^1\mathcal{M}$ can be constructed starting from the above Euclidean process and using the classical stochastic development/parallel transport machinery. Namely, the process $(x_t, v_t)_{t\geq 0}$ in $T^1\mathcal{M}$ is characterised by the fact that the image of $v_t \in T^1_{x_t}\mathcal{M}$ in the fixed unit tangent space $T^1_{x_0}\mathcal{M} \simeq \mathbb{S}^{d-1}$ by the inverse stochastic parallel transport along $(x_s)_{0 \leq s \leq t}$ solves equation (1.1) above.

The isotropic analogue of the process, i.e. the process associated with $\Sigma = \mathrm{Id}$, was introduced in [26], where its was shown, under geometric constraints on the base manifold $\mathcal{M}$, that as the parameter σ goes from zero to infinity, then the sample paths of the process $(x_{\sigma^2 t})_{t\geq 0}$ interpolates in a precise sense between geodesics and Brownian paths. In [1], it was given the name kinetic Brownian motion, and convergence was proved in all generality. The authors also fully determined, for a fixed intensity parameter σ, the Poisson boundary of the process if the base manifold is rotationally invariant. It is easy to prove that, as σ tends to 0, the generalised anisotropic motion converges to the geodesic equations as well; in this work, we intend to describe the limit $\sigma \to \infty$.

The motivation to introduce anisotropy in this context is twofold. From an applied point of view, the kinetic Brownian motion is a simple, yet very reasonable model for the dynamics of a mesoscopic spherical particle with bounded velocity in an isotropic heat bath. Compared to the standard Langevin dynamics where the velocities are Gaussian, the fact that the velocities are here of unit norm is perfectly consistent with special relativity theory. The homogenisation phenomena shown in [1] illustrates the fact that the scaling limit of the process, i.e. the macroscopic behaviour of the particle is nevertheless diffusive, as anticipated. Now, if the geometry of the mesoscopic particle under consideration is not spherical, or if the heat bath is anisotropic, the dynamics of the velocity process has to be anisotropic, see e.g. [19, 11, 24] and the references therein. In that context, the velocity evolution given by the stochastic differential equation (1.1)

with $\Sigma \neq \mathrm{Id}$ is very natural. As we will see below and with this applied point of view, the main result of this article guarantees that the macroscopic behaviour of the particle is still diffusive, with an explicit anisotropy matrix.

From a more theoretical point of view, the introduction of anisotropy is also unavoidable if one wants to generalise the results of [1] to an infinite dimensional setting, say to an infinite dimensional Hilbert space. Indeed, doing so, one quickly faces the problem of defining spherical Brownian motion in this context. Looking at equation (1.1), the orthogonal projection makes perfect sense in a Hilbert setting but we have to give meaning to the driving Brownian motion B. This can naturally be done using the notion of abstract Wiener space, see e.g. [17, 18] or [34, Chapter 8]. Roughly speaking, in that framework the driving process in (1.1) has to belong to the image of a radonifying injection, hence introducing a Hilbert-Schmidt covariance operator. In a finite dimensional setting, the action of this Hilbert-Schmidt operator amounts to replacing the standard Brownian motion B by a Brownian motion with covariance $\Sigma \neq \mathrm{Id}$, i.e. to replace the isotropic noise driving kinetic Brownian motion by an anisotropic noise; this justifies our choice of dynamics for the velocity process.

Our goal in this paper is to exhibit the asymptotics of the time rescaled process $(x_{\sigma^2 t}^\sigma, v_{\sigma^2 t}^\sigma)_{t \geq 0}$ as the intensity parameter σ goes to infinity. More precisely, we show that in both Euclidean and Riemannian contexts, its first projection converges to an anisotropic Brownian motion. The presence of anisotropy drastically complexifies the approach and computations compared to the isotropic framework. Namely, in the isotropic Euclidean setting considered in Section 2.2 of [1] and which is the core of the proof when associated with rough paths techniques, the homogenisation of kinetic Brownian motion was proved using Itô calculus and standard martingale techniques. As it will be clear in Section 2 below, the Doob–Meyer decomposition of the velocity process given by equation (1.1) gets more involved here, its invariant measure is not likely to be easy to describe, and martingale techniques need explicit solutions of the Poisson equation which seems hopeless in this context. In fact, guessing a formula for the invariant measure of the v^σ on the sphere before reading the statement of Proposition 1.1 does not seem obvious.

For this reason, we adopt a different approach and point of view here. Our proof of homogenisation for the time rescaled version of the process $(x_t^\sigma, v_t^\sigma)_{t \geq 0}$ is indeed essentially based on quantitative mixing properties of the velocity process. We show in particular that

Proposition 1.1 (Lemma 2.1 and Proposition 2.4 below). *The process v_t^σ solution of* (1.1) *is ergodic in $\mathbb{S}^{d-1}$ with an explicit invariant measure μ whose density with respect to the uniform measure $\mathrm{d}\theta$ on the sphere is given by*

$$\frac{\mathrm{d}\mu}{\mathrm{d}\theta}(\theta) = \frac{\|A^{-1}\theta\|^{1-d}}{\int_{\mathbb{S}^{d-1}} \|A^{-1}\theta\|^{1-d}\mathrm{d}\theta}.$$

In particular, the invariant measure μ are well as the trajectories are invariant under all the coordinate reflections

$$(\theta^1, \cdots, \theta^i, \cdots, \theta^d) \mapsto (\theta^1, \cdots, -\theta^i, \cdots, \theta^d), \quad 1 \leq i \leq d. \tag{1.3}$$

Moreover, there exists a positive constant τ such that, if $\mathcal{F}_{[a,b]}$ denotes the σ-algebra generated by the unit speed ($\sigma = 1$) velocity process v_t, for $a \leq t < b$, then for any $0 \leq s < t$ and any bounded measurable real-valued random variables P and F that are $\mathcal{F}_{[0,s]}$ and $\mathcal{F}_{[t,\infty]}$-measurable, respectively, we have

$$\left|\mathbb{E}_\mu[PF] - \mathbb{E}_\mu[P]\,\mathbb{E}_\mu[F]\right| \lesssim |P|_\infty |F|_\infty \, \mathrm{e}^{-(t-s)/\tau}. \tag{1.4}$$

The above strong mixing and symmetry properties of the velocity process are the key ingredients to establish the homogenisation of the anisotropic version of kinetic Brownian motion in the Euclidean setting. Indeed, we have the following result.

Theorem A (Theorem 3.6 below). *Let $(x_t^\sigma, v_t^\sigma)_{t\geq 0}$ with values in $T^1\mathbb{R}^d$ be the solution of equation (1.1) and (1.2), starting from (x_0, v_0) where x_0 is fixed and v_0 chosen at random according to μ. Then as σ goes to infinity, the time rescaled process $(x_{\sigma^2 t}^\sigma)_{t\in[0,1]}$ converges in law to a Brownian motion in the Euclidean space $\mathbb{R}^d$, with covariance matrix $\mathrm{diag}(c_1, \cdots, c_d)$ where*

$$c_i := 2\int_0^{+\infty} \mathbb{E}_\mu[v_0^i v_t^i]\mathrm{d}t, \quad 1 \leq i \leq d.$$

Our strategy of proof in the manifold setting consists in establishing that the rough path lift of $(x_{\sigma^2 t}^\sigma)_{t\geq 0}$ converges to the Stratonovich rough path lift of a Brownian motion with the above covariance. To do so, we use again the strong mixing properties of the velocity process, associated with a Lamperti-type criterion to ensure the tightness of the lift in rough path topology — see Lemmas 3.1 and 3.2 below. We then identify the limit process by showing that it has to be a stationary process with independent Gaussian increments on the nilpotent group associated with the rough path structure, see Theorem 3.6.

Using the fact that the notion of stochastic development amounts to solving a stochastic differential equation and that the Itô map is continuous with respect to the rough paths topology, one can conclude that the previous Euclidean statement actually holds in that more general setting. Anisotropic Brownian motion on $\mathcal{M}$ is defined as the stochastic development of an anisotropic Brownian motion in $T_{x_0}\mathcal{M}$.

Theorem B (See section 3.3 below). *Let $(\mathcal{M}, g)$ be a complete and stochastically complete Riemannian manifold and let $(x_t^\sigma, v_t^\sigma)_{t\geq 0}$ be the process with values in $T^1\mathcal{M}$ characterised by the fact that the image of $v_t \in T_{x_t}^1\mathcal{M}$ in the fixed unit tangent space $T_{x_0}^1\mathcal{M} \simeq \mathbb{S}^{d-1}$ by the inverse stochastic parallel transport along $(x_s)_{0\leq s\leq t}$ solves equation (1.1) in $T_{x_0}^1\mathcal{M}$. Then as σ goes to infinity, the time rescaled process $(x_{\sigma^2 t}^\sigma)_{t\in[0,1]}$ converges in law to an anisotropic Brownian motion on the base manifold $\mathcal{M}$.*

The use of rough paths is not a mere commodity. Indeed, on a general manifold, the arguments of the proof of Theorem A will fail, and we have to rely on other methods to make our way up to Riemannian manifolds. As we discuss in more details in example 4.2.1, one can actually prove that for kinetic motions such that property (1.3) does not hold, it is possible that the manifold-valued behaviour be non-Brownian in the limit, a phenomenon not present in the euclidean setting. The rough path theoretic approach provides us with a framework to deal with these subtleties; in fact the limit in the rough equivalent of Theorem A (see Theorem 3.6) is the canonical (Stratonovich) Brownian rough path under condition (1.3), but could be other lifts of the Brownian motion for other kinetic processes.

As will be clear from the proof of Theorem A, the homogenisation phenomenon holds as soon as the mixing properties of the velocity process and the symmetry of the trajectories described in Proposition 1.1 hold. In other words, the conclusion of Theorem B is valid as soon as the process $(x_t^\sigma, v_t^\sigma)_{t\geq 0}$ we consider is the stochastic development of a velocity process satisfying the conclusions of Proposition 1.1. In particular, our proof actually applies even if $(v_t)_{t\geq 0} = (v_t^\sigma)_{t\geq 0}$ is an ergodic Markov process with jumps on $\mathbb{S}^{d-1}$ as soon as the conditions (1.3) and (1.4) are fulfilled.

Theorem C (Theorem 4.1 below). *Let $(\mathcal{M}, g)$ be a complete and stochastically complete Riemannian manifold and let $(x_t^\sigma, v_t^\sigma)_{t\geq 0}$ be the process with values in $T\mathcal{M}$ characterised*

by the fact that the image of $v_t \in T_{x_t}\mathcal{M}$ in the fixed tangent space $T_{x_0}\mathcal{M} \simeq \mathbb{R}^d$ by the inverse stochastic parallel transport along $(x_s)_{0 \le s \le t}$ satisfies the conditions (1.3) and (1.4). Then as σ goes to infinity, the time rescaled process $(x^\sigma_{\sigma^2 t})_{t \in [0,1]}$ converges in law to an anisotropic Brownian motion on the base manifold $\mathcal{M}$.

See Theorem 4.1 for a precise statement. In this level of generality, in Section 4.2 we recover classical results, amongst which Pinsky's so-called random flight [33] and time-dependent variations of it; the anisotropic Langevin diffusion, where v is an anisotropic Ornstein-Uhlenbeck process; and linear interpolation of symmetric random walks as in [8]. It is unclear whether or not the methods of X.M. Li [27, 28] or Herzog, Hottovy and Volpe [20] can get back such a result. In a somewhat independent direction, the interesting work [10] of Chevyrev and coauthors studies this kind of convergence in deterministic systems.

The outline of the article is the following. In the next Section 2, we study the velocity process solution of equation (1.1). We characterise its invariant measure and establish the mixing properties which are the key ingredients in our approach of the homogenisation phenomenon. Section 3 is then devoted to the proofs of our main Theorems A and B. More precisely, in Section 3.1, we show the tightness of the rough path lift of the process in the Euclidean setting. In Section 3.2, we then identify the limit as a Brownian motion on the underlying two-step nilpotent Lie group. This completes the proof of Theorem A in the Euclidean setting. Finally, in Section 3.3, we use the continuity of the Itô map to extend the proof of homogenisation to an arbitrary complete stochastically complete Riemannian manifold. The last section consists in developments, including Theorem C and comments in Section 4.1, and various examples in Section 4.2.

2 Mixing properties of the velocity process

Let $(B_t)_{t \ge 0}$ be a Euclidean Brownian motion in $\mathbb{R}^d$ with non degenerate covariance matrix Σ. Without loss of generality, up to an appropriate choice of coordinate system, we can assume that the matrix Σ is diagonal, with square root A, namely

$$\Sigma = \mathrm{diag}\left(\alpha_1^2, \cdots, \alpha_d^2\right), \quad A = \mathrm{diag}\left(\alpha_1, \cdots, \alpha_d\right).$$

Let us recall that, by definition, the anisotropic velocity process $(v_t) = (v_t^1, \cdots, v_t^d)$ with values in $\mathbb{S}^{d-1} \subset \mathbb{R}^d$ and with intensity $\sigma > 0$ is the solution of the Stratonovich stochastic differential equation

$$\mathrm{d}v_t = \sigma \Pi_{v_t^\perp} \circ \mathrm{d}B_t,$$

where $\Pi_{v_t^\perp}$ denotes the projection on the orthogonal of v_t. Equivalently, there exist a standard Euclidean Brownian motion $(W_t)_{t \ge 0}$ such that v_t satisfies the Itô stochastic differential equation

$$\mathrm{d}v_t = \sigma \Pi_{v_t^\perp} A \mathrm{d}W_t - \frac{\sigma^2}{2}\left(\Sigma + \mathrm{tr}(\Sigma)\mathrm{Id} - 2\langle v_t, \Sigma v_t\rangle \mathrm{Id}\right)v_t \mathrm{d}t,$$

or even more explicitly in Euclidean coordinates, for $1 \le i \le d$

$$\mathrm{d}v_t^i = -\frac{\sigma^2}{2}v_t^i\left[\alpha_i^2 + \sum_{j=1}^d \alpha_j^2 - 2\sum_{j=1}^d \alpha_j^2|v_t^j|^2\right]\mathrm{d}t + \sigma\left(\alpha_i \mathrm{d}W_t^i - v_t^i \sum_{j=1}^d \alpha_j v_t^j \mathrm{d}W_t^j\right). \quad (2.1)$$

In the following, d and Σ are fixed, and we write $f \lesssim g$ for some quantities f and g whenever $f \le Cg$ for a constant $C > 0$ independent of any other parameter.

2.1 Invariant measure

The object of this section is to establish that the velocity process $(v_t)_{t \geq 0}$ is ergodic in $\mathbb{S}^{d-1}$ and to write down its invariant measure explicitly. From equation (2.1), it is not difficult to express the infinitesimal generator L of the process and try to solve the equation $L^* \mu = 0$. Nevertheless, since we are working on the sphere, integrations by parts and computations are quite unpleasant, and we prefer to introduce a natural Euclidean lift of the velocity process. Namely, if $\|\cdot\|$ denotes the standard Euclidean norm, consider the $\mathbb{R}^d$-valued process $(u_t)_{t \geq 0}$ starting from $u_0 \neq 0$ such that $v_0 = u_0/\|u_0\|$, and solution of the stochastic differential equation system

$$\mathrm{d}u_t^i = \frac{\sigma^2}{2} \left(-u_t^i \|u_t\|^2 + \alpha_i^2 u_t^i \right) \mathrm{d}t + \sigma \alpha_i \|u_t\| \mathrm{d}W_t^i, \quad 1 \leq i \leq d.$$

Equivalently, it is the solution to the Stratonovich stochastic differential equation

$$\mathrm{d}u_t = -\frac{\sigma^2}{2} \|u_t\|^2 u_t \, \mathrm{d}t + \sigma \|u_t\| \circ \mathrm{d}B_t.$$

Then, a direct application of Itô's formula shows that the projection $u_t/\|u_t\|$ on $\mathbb{S}^{d-1}$ satisfies equation (2.1). To show that u_t is ergodic and find an explicit expression for its invariant measure, let us now perform the simple linear change of variable $y_t := A^{-1} u_t = \Sigma^{-1/2} u_t$. By Itô's formula we get

$$\mathrm{d}y_t^i = \frac{\sigma^2}{2} \left(-\|Ay_t\|^2 y_t^i + \alpha_i^2 y_t^i \right) \mathrm{d}t + \sigma \|Ay_t\| \mathrm{d}W_t.$$

Setting $V_A(y) := -\log \|Ay\| + \frac{1}{2}\|y\|^2$, the infinitesimal generator L_y of y_t is given by

$$L_y = \frac{\sigma^2}{2} \|Ay\|^2 L_0, \quad \text{where} \quad L_0 := \left(-\nabla V_A \cdot \nabla + \Delta \right).$$

The diffusion process with generator L_0 is naturally ergodic with invariant measure proportional to $\mathbf{e}^{-V_A}$ so that $(y_t)_{t \geq 0}$ is also ergodic with invariant measure

$$\nu(\mathrm{d}y) := C_A \, \|Ay\|^{-1} \mathbf{e}^{-\frac{1}{2}\|y\|^2} \, \mathrm{d}y,$$

where C_A is a normalizing constant. In other words, the Euclidean lift $(u_t)_{t \geq 0}$ of $(v_t)_{t \geq 0}$ is ergodic in $\mathbb{R}^d$ and its invariant measure is proportional to $\|\cdot\|^{-1}$ times the centred Gaussian measure with covariance Σ. One can then compute the invariant measure of the velocity process as the image measure of the latter with respect to the projection on the sphere.

Lemma 2.1. *The velocity process $(v_t)_{t \geq 0}$ is ergodic in $\mathbb{S}^{d-1}$ and its invariant measure μ is absolutely continuous with respect to the uniform measure $\mathrm{d}\theta$ on the sphere, with a density given by*

$$\frac{\mathrm{d}\mu}{\mathrm{d}\theta}(\theta) = \frac{\|A^{-1}\theta\|^{1-d}}{\int_{\mathbb{S}^{d-1}} \|A^{-1}\theta\|^{1-d} \mathrm{d}\theta}.$$

In particular, the invariant measure μ of the velocity process is invariant under all the coordinate reflections $(\theta_1, \cdots, \theta_i, \cdots, \theta_d) \mapsto (\theta_1, \cdots, -\theta_i, \cdots, \theta_d)$, for $1 \leq i \leq d$.

Proof. For any bounded measurable test function f on $\mathbb{S}^{d-1}$, we have

$$
\begin{aligned}
\int_{\mathbb{S}^{d-1}} f(v)\mu(\mathrm{d}v) &= C_A \int_{\mathbb{R}^d} f\left(\frac{Ay}{\|Ay\|}\right) \frac{\mathrm{e}^{-\frac{1}{2}\|y\|^2}}{\|Ay\|}\mathrm{d}y \\
&= C_A \int_{\mathbb{R}^d} f\left(\frac{u}{\|u\|}\right) \|u\|^{-1}\mathrm{e}^{-\frac{1}{2}\|A^{-1}u\|^2}\frac{\mathrm{d}u}{\det A} \\
&= C_A' \int_0^{+\infty} \int_{\mathbb{S}^{d-1}} f(\theta) r^{-1}\mathrm{e}^{-\frac{1}{2}r^2\|A^{-1}\theta\|^2}r^{d-1}\mathrm{d}r\mathrm{d}\theta \\
&= \frac{\int_{\mathbb{S}^{d-1}} f(\theta)\|A^{-1}\theta\|^{1-d}\mathrm{d}\theta}{\int_{\mathbb{S}^{d-1}} \|A^{-1}\theta\|^{1-d}\mathrm{d}\theta}. \quad \square
\end{aligned}
$$

The next figures illustrate the relation between the covariance matrix Σ, the sample paths of the velocity process (v_t) and its invariant measure μ. The colour map on the sphere is chosen according to the value of the density of the invariant measure: small values of $\|A^{-1}\theta\|^{1-d}$ are represented in light grey whereas large values are represented in dark grey.
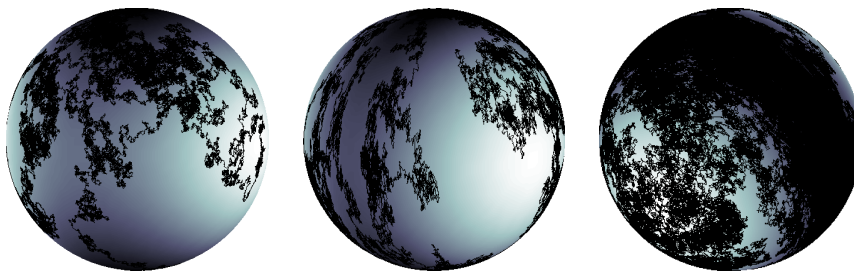

Figure 2: From left to right, sample paths of the velocity process and colour map of the invariant probability measure for $\Sigma = \mathrm{diag}(1, 1.1, 1.2)$, $\Sigma = \mathrm{diag}(1, 4, 9)$, and $\Sigma = \mathrm{diag}(1, 100, 100)$.

Remark 2.2. Let us emphasise here that the invariant measure μ of the velocity process actually differs from the projected Gaussian measure with covariance Σ, also known as angular Gaussian distribution, which, at first sight, could seem like a natural candidate for the velocity's equilibrium measure. Namely, if f is a bounded measurable test function on the sphere, and if X is a Gaussian variable in $\mathbb{R}^d$ with law $\mathcal{N}(0, \Sigma)$, we have indeed

$$
\mathbb{E}\left[f\left(\frac{X}{\|X\|}\right)\right] = \frac{\int_{\mathbb{S}^{d-1}} f(\theta)\|A^{-1}\theta\|^{-d}d\theta}{\int_{\mathbb{S}^{d-1}} \|A^{-1}\theta\|^{-d}d\theta}.
$$

In other words, the invariant measure μ admits a density proportional to $\|A^{-1}\theta\|$ with respect to the standard projected Gaussian measure of covariance Σ.

Remark 2.3. Going back to the modelisation point of view mentioned in the introduction, where $(v_t)_{t\geq 0}$ is thought as the velocity of a mesoscopic particle in an anisotropic heat bath, the invariant measure μ also differs from the standard choices for equilibrium measure in directional statistics, such as the Von Mises–Fisher distribution, Fisher–Bingham distribution or wrapped Brownian distributions, see Sections 9.3 and 9.4 of [32] and the references therein. We emphasise here the fact that the dynamics governed by equation (1.1) is fully intrinsic so that the measure μ is a simple and natural candidate

to model anisotropic data; it also has natural interpretation in terms of projection of the invariant measure of the Euclidean lift $(u_t)_{t \geq 0}$.

2.2 Mixing properties

Let us now establish the strong mixing properties of the velocity process that will be our main tool in the proof of the homogenisation result, Theorem A. To avoid changes in the time scale, we fix $\sigma = 1$, from here to the end of the section. We also introduce a few additional notations. If λ is a probability distribution on $\mathbb{S}^{d-1}$, let $\mathbb{P}_\lambda$ be a probability measure under which the velocity $(v_t)_{t \geq 0}$ solves equation (1.1) with initial condition $v_0 \sim \lambda$, and $\mathbb{E}_\lambda$ its associated expectation. We denote by $(P_t)_{t \geq 0}$ the semigroup associated to v, acting on continuous functions $f : \mathbb{S}^{d-1} \to \mathbb{R}$, and by $(P_t^*)_{t \geq 0}$ its dual, acting on probability measures on $\mathbb{S}^{d-1}$. In other words,

$$P_t f(x) := \mathbb{E}_{\delta_x}[f(v_t)] \qquad \text{and} \qquad P_t^* \lambda := \mathcal{L}(v_t | v_0 \sim \lambda),$$

for any such f and λ.

To get to the second part of Proposition 1.1, we use the well-known fact that since the velocity process $(v_t)_{t \geq 0}$ is an elliptic diffusion in a compact Riemannian manifold, here the unit sphere, with invariant probability measure μ, we have the estimate

$$\|P_t^* \lambda - \mu\|_{\mathrm{TV}} \lesssim \exp(-t/\tau) \tag{2.2}$$

for any probability λ on $\mathbb{S}^{d-1}$, for some positive constant τ. Given an interval $[a, b)$ of $[0, \infty)$, define $\mathcal{F}_{[a,b)}$ as the σ-algebra generated by the unit speed velocity process v_t, for $a \leq t < b$. We write $A \in \mathcal{F}_{[a,b)}$ to say that a random variable is $\mathcal{F}_{[a,b)}$-measurable.

Proposition 2.4. *For any $0 \leq s < t$ and any bounded measurable real-valued random variables $P \in \mathcal{F}_{[0,s)}$ and $F \in \mathcal{F}_{[t,\infty)}$, we have*

$$\left| \mathbb{E}_\mu[PF] - \mathbb{E}_\mu[P] \, \mathbb{E}_\mu[F] \right| \lesssim |P|_\infty |F|_\infty \, \mathbf{e}^{-(t-s)/\tau}.$$

Proof. Since

$$\left| \mathbb{E}_\mu[PF] - \mathbb{E}_\mu[P]\mathbb{E}_\mu[F] \right| \leq |P|_\infty \mathbb{E}_\mu\left[\left| \mathbb{E}_\mu[F | \mathcal{F}_{[0,s]}] - \mathbb{E}_\mu[F] \right| \right],$$

by the Markov property, it suffices to prove that one has

$$\left| \mathbb{E}_{P_u^* \lambda}[G] - \mathbb{E}_\mu[G] \right| \lesssim |G|_\infty \mathbf{e}^{-u/\tau}, \tag{2.3}$$

for any probability measure λ on $\mathbb{S}^{d-1}$ and any real-valued measurable functional G. By a monotone class argument, it suffices to prove estimate (2.3) for elementary functionals of the form $G = g(v_{t_1}, \ldots, v_{t_k})$, for some bounded continuous real-valued function g on $(\mathbb{R}^d)^k$ and times $t_1 \leq \cdots \leq t_k$. But since the diffusion has the Feller property, the function $\overline{g}(v_0) := \mathbb{E}_{v_0}\big[g(v_{t_1}, \ldots, v_{t_k})\big]$ is continuous on the sphere, so we get (2.3) in that case by applying (2.2) to $\overline{g}$. $\qquad\square$

The remainder of the section is devoted to the proof of the technical Lemma 2.6, that states an estimate about iterated integrals involving the covariances between the coordinates of the unit speed velocity process. Given a collection of positive times $s_1, \ldots, s_n$, set $\Delta := \max_{1 \leq k < n}(s_k \wedge s_{k+1})$. We denote by $k_0 \in [\![1, n-1]\!]$ an index where this maximum is attained.

Proposition 2.5. *Under $\mathbb{P} = \mathbb{P}_\mu$, and for any indices $1 \leq j_1, \ldots, j_n \leq d$ and times $s_1, \cdots, s_n \geq 0$,*

$$\left| \mathbb{E}\big[v_{s_1}^{j_1} \cdots v_{s_1 + \cdots + s_n}^{j_n} \big] \right| \lesssim \mathbf{e}^{-\Delta/\tau}.$$

Proof. For $1 \leq i \leq n$, set $t_i := s_1 + \cdots + s_i$, and define the bounded quantities

$$V_- := v_{t_1}^{j_1} \cdots v_{t_{k_0-1}}^{j_{k_0}-1}, \qquad V_0 := v_{t_{k_0}}^{j_{k_0}}, \qquad V_+ := v_{t_{k_0+1}}^{j_{k_0+1}} \cdots v_{t_n}^{j_n}.$$

Note that V_0 is centred. Applying Proposition 2.4 twice, this decomposition gives

$$\begin{aligned}
\left| \mathbb{E}\left[v_{s_1}^{j_1} \cdots v_{s_1+\cdots+s_n}^{j_n} \right] \right| &= \left| \mathbb{E}[V_- V_0 V_+] - \mathbb{E}[V_-]\mathbb{E}[V_0]\mathbb{E}[V_+] \right| \\
&\leq \left| \mathbb{E}[V_- V_0 V_+] - \mathbb{E}[V_-]\mathbb{E}[V_0 V_+] \right| + |V_-|_\infty \left| \mathbb{E}[V_0 V_+] - \mathbb{E}[V_0]\mathbb{E}[V_+] \right| \\
&\lesssim |V_-|_\infty |V_0 V_+|_\infty \mathbf{e}^{-s_{k_0}/\tau} + |V_-|_\infty |V_0|_\infty |V_+|_\infty \mathbf{e}^{-s_{k_0+1}/\tau} \\
&\lesssim \mathbf{e}^{-\Delta/\tau}. \hspace{4cm} \square
\end{aligned}$$

Lemma 2.6. *Suppose $\mathbb{P} = \mathbb{P}_\mu$. Given a positive integer n, we have*

$$\int_{0 \leq t_1 \leq \cdots \leq t_{2n} \leq T} \left| \mathbb{E}_\mu\left[v_{t_1}^{i_1} \cdots v_{t_{2n}}^{i_{2n}} \right] \right| \, \mathrm{d}t_1 \ldots \mathrm{d}t_{2n} \lesssim_n T^n$$

for any indices $1 \leq i_1, \cdots, i_{2n} \leq d$.

Proof. The idea is to apply Proposition 2.5 with the largest Δ possible for each tuple $(t_1, \cdots, t_{2n})$. Write first

$$\int_{0 \leq t_1 \leq \cdots \leq t_{2n-1}} \left| \mathbb{E}\left[v_{t_1}^{i_1} \cdots v_{t_{2n}}^{i_{2n}} \right] \right| \, \mathrm{d}t_1 \ldots \mathrm{d}t_{2n} \leq \int_{[0,T]^{2n}} \left| \mathbb{E}\left[v_{s_1}^{i_1} \cdots v_{s_1+\cdots+s_{2n}}^{i_{2n}} \right] \right| \, \mathrm{d}s_1 \ldots \mathrm{d}s_{2n}.$$

Fix now the tuple $(s_1, \cdots, s_{2n})$, and set

$$\Delta(s) := \max_{1 \leq k < 2n} \left(s_k \wedge s_{k+1} \right),$$

so the integrand in the right hand side above is bounded above by a constant multiple of $\mathbf{e}^{-\Delta(s)/\tau}$, from Proposition 2.5.

The rest is combinatorics. We first sort the indices k of the gaps s_k according to the value of s_k with respect to $\Delta = \Delta(s)$. Set $a := \min\{k \in [\![1, 2n]\!] : s_k = \Delta\}$. Then, note that there are at most n gaps s_k of size larger than Δ: otherwise, two of them would be consecutive, and Δ would not be optimal. This is the same as saying that there are at least n small gaps $s_k \leq \Delta$, including s_a. Define $1 \leq b_1 < \cdots < b_{n-1} \leq 2n$ as the first $(n-1)$ indices different from a corresponding to gaps of size at most Δ. In other words, if $s_k \leq \Delta$, then either $k = b_i$ for some $1 \leq i < n$, $k = a$, or $k > a, b_{n-1}$. Finally, denote by $1 \leq c_1 < \cdots < c_n \leq 2n$ the other indices, so that we have a partition of $\{1, \cdots, 2n\}$ in three sets $A(s) := \{a\}$, $B(s) := \{b_1, \cdots, b_{n-1}\}$ and $C(s) := \{c_1, \cdots, c_n\}$ of fixed sizes. Now, given a fixed partition (α, β, γ) of $[\![1, 2n]\!]$ with $\alpha = \{\alpha_0\}$ of size 1, and the set $\beta = \{\beta_1, \ldots, \beta_{n-1}\}$ of size $n - 1$, we have

$$\begin{aligned}
\left| \mathbb{E}\left[v_{s_1}^{i_1} \cdots v_{s_1+\cdots+s_{2n}}^{i_{2n}} \right] \right| \mathbf{1}_{(A(s),B(s),C(s))=(\alpha,\beta,\gamma)} &\lesssim \mathbf{e}^{-\Delta(s)/\tau} \mathbf{1}_{s_{\beta_1},\ldots,s_{\beta_{n-1}} \leq \Delta(s)} \\
&\lesssim \mathbf{e}^{-s_{\alpha_0}/\tau} \mathbf{1}_{s_{\beta_1},\ldots,s_{\beta_{n-1}} \leq s_{\alpha_0}},
\end{aligned}$$

from which we get

$$\int_{[0,T]^{2n}} \left| \mathbb{E}\left[v_{s_1}^{i_1} \cdots v_{s_1+\cdots+s_{2n}}^{i_{2n}} \right] \right| \mathbf{1}_{(A(s),B(s),C(s))=(\alpha,\beta,\gamma)} \, \mathrm{d}s_1 \ldots \mathrm{d}s_{2n}$$

$$\lesssim T^n \int_0^T \mathbf{e}^{-s/\tau} s^{n-1} \mathrm{d}s \lesssim_n T^n$$

and the result of the lemma, by summing over the set of all partitions (α, β, γ) of $[\![1, 2n]\!]$ with the above size. $\hspace{3cm} \square$

3   Proof of the main result

Let us now describe how the mixing properties of the velocity process derived in Section 2.2 imply the homogenisation for the time rescaled position process $(x^\sigma_{\sigma^2 t})_{t\geq 0}$, as σ goes to infinity, in both Euclidean and Riemannian framework. As mentioned in the introduction, we will actually work with a rough path lift of the kinetic process. We refer the reader to [15, 5] for gentle introductions to rough paths theory, and given $\gamma \in (0,1)$, we denote by $\mathsf{RP}(\gamma) = \mathsf{RP}^\gamma([0,1], \mathbb{R}^d)$ the set of weak geometric γ-Hölder rough paths.

The strategy is a classical Prokhorov argument: we first show that the family of laws of the rescaled processes is tight, then identify a unique candidate for the limit.

Notations   We are interested in the stationary case $\mathbb{P} := \mathbb{P}_\mu$, where μ is the invariant measure of the velocity, as described in Lemma 2.1. Define $X^\sigma : t \mapsto x^\sigma_{\sigma^2 t}$, so that we are interested in the limiting behaviour of $(X^\sigma_t)_{t\geq 0}$. To make good use of the mixing properties of v such as Proposition 2.4 without having to change the time scale, from now on $(v_t)_{t\geq 0}$ *will always stand for* $(v^\sigma_t)_{t\geq 0}$ *with* $\sigma = 1$. With this convention, we can express the increments of X^σ as

$$X_t - X_s = \frac{1}{\sigma^2} \int_{\sigma^4 s}^{\sigma^4 t} v_u \mathrm{d}u.$$

The process X^σ being $\mathcal{C}^1$, it admits a canonical rough path lift $\mathbf{X}^\sigma = (X^\sigma, \mathbb{X}^\sigma)$, where $\mathbb{X}^\sigma$ is defined by

$$\mathbb{X}^\sigma_{ts} := \int_s^t (X^\sigma_u - X^\sigma_s) \otimes \mathrm{d}X^\sigma_u = \frac{1}{\sigma^4} \int_{\sigma^4 s}^{\sigma^4 t} \int_{\sigma^4 s}^{u} v_z \otimes v_u \mathrm{d}z\, \mathrm{d}u.$$

Our proof relies on the algebraic properties of rough paths. Namely, that in the 2-step nilpotent group $G \subset \mathbb{R} \oplus \mathbb{R}^d \oplus (\mathbb{R}^d)^{\otimes 2}$, the process $\mathbf{x}^\sigma : t \mapsto (1, X^\sigma_t, \mathbb{X}^\sigma_{t0})$ has increments

$$(\mathbf{x}^\sigma_s)^{-1} \mathbf{x}^\sigma_t = (1, X^\sigma_t - X^\sigma_s, \mathbb{X}^\sigma_{ts}),$$

which, using the above expressions, are measurable with respect to $\sigma\big((v_u)_{\sigma^4 s \leq u < \sigma^4 t}\big)$.

Recall that we write $f \lesssim g$ for some quantities f and g when there exists a positive constant $C > 0$ depending on Σ alone such that $f \leq Cg$. If C is allowed to depend on a parameter, say p, we write $f \lesssim_p g$.

3.1   Tightness in rough paths space

We first establish that the family of processes (X^σ_t) and their rough paths lifts are tight for the corresponding topology. To do so, we use a standard Lamperti criterion, namely we have the following lemma.

Lemma 3.1. *For every* $a \geq 1$,

$$\sup_{\sigma > 0} \mathbb{E}\big[|X^\sigma_t - X^\sigma_s|^a\big] \lesssim_a |t - s|^{a/2}.$$

Proof. Given any positive time T and any positive integer n, we show that one has

$$\mathbb{E}\left[\left|\int_0^T v_t\, \mathrm{d}t\right|^{2n}\right] \leq C_n T^n \tag{3.1}$$

for some positive constant C_n depending only on n. The inequality of the lemma follows as a consequence since for any positive integer n such that $2n \geq a$, we have

$$\mathbb{E}\big[|X^\sigma_t - X^\sigma_s|^a\big] = \mathbb{E}\big[|X^\sigma_{t-s}|^a\big] \leq \frac{1}{\sigma^{2a}} \mathbb{E}\left[\left|\int_0^{\sigma^4(t-s)} v_u\, \mathrm{d}u\right|^{2n}\right]^{a/2n} \leq C_n^{a/2n}(t-s)^{a/2}.$$

Given $T > 0$ and $n \in \mathbb{N}^*$, we have

$$\mathbb{E}\left[\left|\int_0^T v_t \, dt\right|^{2n}\right] = \mathbb{E}\left[\left(\sum_{1 \le i \le d}\left(\int_0^T v_t^i \, dt\right)^2\right)^n\right]$$

$$= \sum_{1 \le i_1, \cdots, i_n \le d} \int_{[0,T]^{2n}} \mathbb{E}\left[v_{t_1}^{i_1} v_{t_2}^{i_1} \cdots v_{t_{2n-1}}^{i_n} v_{t_{2n}}^{i_n}\right] dt_1 \cdots dt_{2n},$$

with the following estimate for each individual term on the right hand side. Fix $1 \le j_k \le d$, for $1 \le k \le 2n$. For any permutation $\phi \in \mathfrak{S}_{2n}$, we have from Lemma 2.6

$$\int_{[0,T]^{2n}} \mathbb{E}\left[v_{t_1}^{j_1} \cdots v_{t_{2n}}^{j_{2n}}\right] \mathbf{1}_{t_{\phi(1)} < \cdots < t_{\phi(2n)}} \, dt = \int_{0 \le t_1 \le \cdots \le t_{2n} \le T} \mathbb{E}\left[v_{t_1}^{j_{\phi(1)}} \cdots v_{t_{2n}}^{j_{\phi(2n)}}\right] dt_1 \cdots dt_{2n}$$

$$\lesssim_n T^n,$$

from which the result of the Lemma follows by summation over ϕ and j. $\qquad \square$

We use the Hilbert-Schmidt norm $|\cdot|$ on $\mathbb{R}^d \otimes \mathbb{R}^d \simeq L(\mathbb{R}^d) \simeq \mathbb{R}^{d^2}$; it coincides with the Euclidean norm on $\mathbb{R}^{d^2}$.

Lemma 3.2. *For every $a > 0$,*

$$\sup_{\sigma > 0} \mathbb{E}\left[|\mathbb{X}_{ts}^{\sigma}|^a\right] \lesssim_a |t - s|^a.$$

Proof. As above, the inequality of the statement follows from an inequality of the form

$$\mathbb{E}\left[\left|\int_{0 \le s \le t \le T} v_s \otimes v_t \, ds dt\right|^{2n}\right] \le C_n T^{2n},$$

for some positive constant C_n depending only on n. Fix $T > 0$ and $n \in \mathbb{N}^*$, and set for $\ell \in [\![1, d]\!]^{4n}$

$$I_\ell := \int_{0 \le s_1 \le t_1 \le T} \cdots \int_{0 \le s_{2n} \le t_{2n} \le T} \mathbb{E}\left[v_{t_1}^{\ell_1} v_{s_1}^{\ell_2} \cdots v_{t_{2n}}^{\ell_{4n-1}} v_{s_{2n}}^{\ell_{4n}}\right] ds_1 dt_1 \cdots ds_{2n} dt_{2n},$$

so we have

$$\mathbb{E}\left[\left|\int_0^T \int_0^t v_s \otimes v_t \, ds dt\right|^{2n}\right] = \mathbb{E}\left[\left(\sum_{1 \le i,j \le d}\left(\int_0^T \int_0^t v_s^i v_t^j \, ds dt\right)^2\right)^n\right]$$

$$= \sum_{i,j \in [\![1,d]\!]^n} I_{i*j}$$

with $i * j = (i_1, j_1, i_1, j_1, \cdots, i_k, j_k, i_k, j_k)$. As in Lemma 3.1, estimating each I_{i*j} using Lemma 2.6 does the job. $\qquad \square$

One can then apply the Kolmogorov-Lamperti tightness criterion for rough paths stated in Corollary A.12 of [16] to get the following result from Lemma 3.1 and Lemma 3.2.

Corollary 3.3. *Pick $\gamma < 1/2$. The family $\{\mathcal{L}(\mathbf{X}^\sigma)\}_{\sigma > 0}$ of distributions on $\mathrm{RP}(\gamma)$ is tight.*

Remark 3.4. Note that the usual rough path space $\mathrm{RP}(\gamma)$ actually involves more iterated integrals as $\gamma \to 0$. For instance, one has to consider the quantity

$$\int_s^t \int_s^u (X_v - X_s) \otimes dX_v \otimes dX_u$$

for rough differential equations to be well-posed whenever $1/4 < \gamma \le 1/3$. However, there is a canonical continuous injection from $\mathrm{RP}(\gamma)$ to $\mathrm{RP}(\gamma')$ for $0 < \gamma' < \gamma$, constructed in much the same way as the Riemann integral from the limit Riemann sum. Accordingly, the convergence in $\mathrm{RP}(\gamma)$, $\gamma > 1/3$ is sufficient to prove it for smaller γ'; this is what we do in what follows.

Remark also that, for $\gamma > 1/2$, we have $\mathrm{RP}(\gamma) \simeq \mathcal{C}^\gamma([0,1], \mathbb{R}^d)$, from which the existence of canonical lift of $\mathcal{C}^\gamma$ paths follows in this range of regularity.

3.2 Brownian limit

The family of processes (X_t^σ) and their lifts being tight for the rough paths topology, in order to establish its convergence, we are left to identify the possible limit process. Our strategy here is to prove that the latter is necessarily a stationary process with independent Gaussian increments on the underlying nilpotent group, and therefore is a Brownian motion. Let us set

$$c_i := 2 \int_0^\infty \mathbb{E}[v_0^i v_t^i] \, \mathrm{d}t.$$

Proposition 3.5. *For every $\gamma < 1/2$, the processes X^σ converge in law in $\mathcal{C}^\gamma([0;1], \mathbb{R}^d)$ to the Brownian motion on $\mathbb{R}^d$ with covariance matrix $\mathrm{diag}(c_1, \cdots, c_d)$, as σ goes to ∞.*

Proof. Stationarity and independence. We first show that any $\mathbb{R}^d$-valued process X whose law $\widehat{\mathbb{P}}$ is a limit point of $(\mathcal{L}(X^\sigma))_{\sigma > 0}$ in $\mathcal{C}^\gamma([0;1], \mathbb{R}^d)$ as σ tends to ∞ has stationary independent increments.

Indeed, since v_0 has distribution the invariant measure of the diffusion v, the increments of X^σ are stationary for every σ, so the increments of X are stationary as well. Fix now $0 \le s_1 < t_1 \le \cdots \le s_n < t_n \le 1$, and bounded continuous functions $F_i : \mathbb{R}^d \to \mathbb{R}$, for $1 \le i \le n$. Fix $\varepsilon > 0$ small enough. From a repetitive use of Proposition 2.4, as used in Proposition 2.5, we have

$$\left| \mathbb{E}\left[\prod_{1 \le i \le n} F_i(X_{t_i-\varepsilon}^\sigma - X_{s_i}^\sigma) \right] - \prod_{1 \le i \le n} \mathbb{E}\left[F_i(X_{t_i-\varepsilon}^\sigma - X_{s_i}^\sigma) \right] \right| \lesssim_n |F_1|_{L^\infty} \cdots |F_n|_{L^\infty} \, \mathbf{e}^{-\sigma^4 \varepsilon/\tau}$$

for some positive constant τ, and we see that

$$\widehat{\mathbb{E}}\left[\prod_{1 \le i \le n} F_i(X_{t_i-\varepsilon} - X_{s_i}) \right] = \prod_{1 \le i \le n} \widehat{\mathbb{E}}\left[F_i(X_{t_i-\varepsilon} - X_{s_i}) \right],$$

sending σ to ∞ along a proper subsequence. Using the boundedness and continuity of the functions F_i and the continuity of the process X, we can send ε to 0 and see that X has independent increments. So X is a Brownian motion; it has null mean since every X_1^σ has null mean, and its covariance is given by the limit of the covariances of the X_1^σ.

Covariance formula. First, it follows from the identity

$$\mathcal{L}(v^1, \cdots, v^i, \cdots, v^n) = \mathcal{L}(v^1, \cdots, -v^i, \cdots, v^n)$$

that different components of X_1 have null covariance since this is the case for different components of X_1^σ. Now, for $1 \le i \le d$, we have

$$\mathbb{E}\left[((X_1^\sigma)^i)^2 \right] = \frac{1}{\sigma^4} \int_0^{\sigma^4} \int_0^{\sigma^4} \mathbb{E}[v_s^i v_t^i] \, \mathrm{d}s \, \mathrm{d}t = \frac{2}{\sigma^4} \int_0^{\sigma^4} \int_t^{\sigma^4} \mathbb{E}[v_s^i v_t^i] \, \mathrm{d}s \, \mathrm{d}t$$

$$= \frac{2}{\sigma^4} \int_0^\infty \int_0^\infty \mathbf{1}_{t+u \le \sigma^4} \mathbb{E}[v_{t+u}^i v_t^i] \mathrm{d}u \, \mathrm{d}t = 2 \int_0^\infty \left(1 - \frac{u}{\sigma^4} \right)_+ \mathbb{E}[v_u^i v_0^i] \mathrm{d}u$$

with $(\cdot)_+$ the positive part. According to Proposition 2.4, the integrand is smaller than a constant multiple of $\exp(-u/\tau)$, uniformly on σ. It is integrable, so we see from Lebesgue dominated convergence theorem that the above variance tends to γ_i. □

Theorem 3.6 (Main Theorem A). *Pick $\gamma < 1/2$. The processes $\mathbf{X}^\sigma$ converge in law in $\mathsf{RP}(\gamma)$ to the Brownian rough path on $\mathbb{R}^d$ with covariance matrix $\mathrm{diag}(\gamma_1, \cdots, \gamma_d)$, as σ goes to ∞.*

Here, the rough path lift of the Brownian motion is understood in the Stratonovich sense: for W a Brownian motion with fixed covariance (or any semimartingale for that matter), it is defined as $\mathbf{w} = (W, \mathbb{W})$ with $\mathbb{W}_{st} := \int_s^t W_u \otimes \mathrm{d}W_u$, where the integral is the Stratonovich one.

Proof. G-valued Lévy process. As above, we first notice that any limit measure of the laws of $(\mathbf{X}^\sigma)_{\sigma > 0}$ turns the canonical process on $\mathsf{RP}(\gamma)$ into a random process with stationary independent increments, in the free nilpotent Lie group of step 2, as a consequence of the corresponding property for $\mathbf{X}^\sigma$. The canonical process on the free nilpotent Lie group of step 2 is thus a continuous Lévy process under any limit law, so, according to Hunt's theorem, we can identify the former from its generator. More specifically, the law of such a process Y is characterised by the action

$$f \mapsto \lim_{t \to 0} \frac{1}{t} \mathbb{E}_e[f(Y_t) - f(e)] \in \mathbb{R}$$

of its generator on smooth functions $f : G \to \mathbb{R}$ with compact support, where e is the unit of G; see [2, Theorem 5.3.3] or [29, Theorem 1.1].

Generator. Let $\widehat{\mathbb{P}}$ be any limit point of the laws of $\mathbf{X}^\sigma$ on $\mathsf{RP}(\gamma)$, and denote by $\mathbf{X} = (X, \mathbb{X})$ its canonical variable. We know from Proposition 3.5 that X is a Brownian motion W; denote by $\mathbf{W} = (W, \mathbb{W})$ its canonical Stratonovich rough path lift, also defined on the space $(\mathsf{RP}(\gamma), \widehat{\mathbb{P}})$. Since the velocity process $v = (v^1, \ldots, v^d)$ and $(v^1, \ldots, v^{i-1}, -v^i, v^{i+1}, \ldots, v^d)$ have the same law for every $1 \le i \le d$, for v_0 distributed according to the invariant measure μ, the antisymmetric part $\mathbb{A}^{\mathbf{X}}_{ts} := \frac{1}{2}(\mathbb{X}_{ts} - {}^t\mathbb{X}_{ts})$ is centred for any $0 \le s \le t \le 1$. We also know from the uniform estimates proved in Lemmas 3.1 and 3.2 that

$$\widehat{\mathbb{E}}\big[|X_t|^2\big] \lesssim t, \quad \widehat{\mathbb{E}}\big[|\mathbb{A}^{\mathbf{X}}_{t0}|^2\big] \lesssim \widehat{\mathbb{E}}\big[|\mathbb{X}_{t0}|^2\big] \lesssim t^2, \tag{3.2}$$

uniformly in $t \in [0, 1]$.

A last piece of notation. Since the set of antisymmetric matrices lies in the tangent space to the free nilpotent Lie group G of step 2, at any point $\mathbf{z} \in G$, any smooth real-valued function f defined on G, with compact support, has a well-defined partial differential $\partial_{\mathbb{A}} f(\mathbf{z})$ in the direction of antisymmetric matrices, defined by the identity

$$\partial_{\mathbb{A}} f(\mathbf{z})(\mathbb{A}) = \frac{\mathrm{d}}{\mathrm{d}t}_{|t=0} f\big(\mathbf{z} + t(0, 0, \mathbb{A})\big),$$

for any $\mathbf{z} = (1, Z, \mathbb{Z}) \in G$ and any antisymmetric matrix $\mathbb{A}$. Setting $\overline{\mathbf{z}} := \big(1, Z, \frac{1}{2}(\mathbb{Z} + {}^t\mathbb{Z})\big)$, we further have

$$\big|f(\mathbf{z}) - f(\overline{\mathbf{z}}) - (\partial_{\mathbb{A}} f)(\overline{\mathbf{z}})(\mathbb{A}^{\mathbf{z}})\big| \lesssim_f |\mathbb{A}^{\mathbf{z}}|^2,$$

since f has compact support. Denote by e the unit of the group G. Denote by $\mathbb{A}^{\mathbf{W}}$ the antisymmetric part of $\mathbb{W}$ and set $\overline{\mathbf{X}}_t := \big(1, X_t, \frac{1}{2}X_t^{\otimes 2}\big) \in G$, so that $\mathbf{X}_t = \overline{\mathbf{X}}_t + (0, 0, \mathbb{A}^{\mathbf{X}})$

and $\mathbf{W}_t = \overline{\mathbf{X}}_t + (0, 0, \mathbb{A}^{\mathbf{W}})$. We have, for some fixed f smooth with compact support,

$$
\left| \frac{1}{t} \widehat{\mathbb{E}} \big[f(\mathbf{X}_t) - f(e) \big] - \frac{1}{t} \widehat{\mathbb{E}} \big[f(\mathbf{W}_t) - f(e) \big] \right|
$$
$$
= \frac{1}{t} \left| \widehat{\mathbb{E}} \Big[f\big(\overline{\mathbf{X}}_t + (0, 0, \mathbb{A}^{\mathbf{X}}_t)\big) - f\big(\overline{\mathbf{X}}_t + (0, 0, \mathbb{A}^{\mathbf{W}}_t)\big) \Big] \right|
$$
$$
\lesssim_f \frac{1}{t} \left| \widehat{\mathbb{E}} \Big[\big((\partial_{\mathbb{A}} f)(\overline{\mathbf{X}}_t) - (\partial_{\mathbb{A}} f)(e) \big)(\mathbb{A}^{\mathbf{X}}_t - \mathbb{A}^{\mathbf{W}}_t) \Big] \right| + \frac{1}{t} \left| \widehat{\mathbb{E}} \big[(\partial_{\mathbb{A}} f)(e)(\mathbb{A}^{\mathbf{X}}_t - \mathbb{A}^{\mathbf{W}}_t) \big] \right|
$$
$$
+ \frac{1}{t} \Big(\widehat{\mathbb{E}} \big[|\mathbb{A}^{\mathbf{X}}_t|^2 \big] + \widehat{\mathbb{E}} \big[|\mathbb{A}^{\mathbf{W}}_t|^2 \big] \Big)
$$
$$
\lesssim_f (1) + (2) + (3).
$$

We show that each term vanishes as t goes to 0, which implies that the two Markov processes $\mathbf{X}$ and $\mathbf{W}$ have the same generator, hence the same distribution. We have first from estimates (3.2) the upper bound

$$
(1) \leq \frac{1}{2t} \widehat{\mathbb{E}} \Big[\sqrt{t} \, \big\| (\partial_{\mathbb{A}} f)(\overline{\mathbf{X}}_t) - (\partial_{\mathbb{A}} f)(e) \big\|^2 \Big] + \frac{1}{2t} \widehat{\mathbb{E}} \Big[\frac{1}{\sqrt{t}} | \mathbb{A}^{\mathbf{X}}_t - \mathbb{A}^{\mathbf{W}}_t |^2 \Big]
$$
$$
\lesssim_f \frac{1}{\sqrt{t}} \widehat{\mathbb{E}} \big[|\overline{\mathbf{X}}_t - e|^2 \big] + \frac{1}{t\sqrt{t}} \widehat{\mathbb{E}} \Big[| \mathbb{A}^{\mathbf{X}}_t |^2 \Big] + \frac{1}{t\sqrt{t}} \widehat{\mathbb{E}} \Big[| \mathbb{A}^{\mathbf{W}}_t |^2 \Big]
$$
$$
\lesssim_f \sqrt{t}.
$$

We also have $(2) = 0$, since $\mathbb{A}^{\mathbf{X}}_t$ and $\mathbb{A}^{\mathbf{W}}_t$ are centred and $\partial_{\mathbb{A}} f(e)$ is linear. Finally, we have $(3) \lesssim t$ from the upper bounds (3.2). We thus have the upper bound

$$
\left| \frac{1}{t} \widehat{\mathbb{E}} \big[f(\mathbf{X}_t) - f(e) \big] - \frac{1}{t} \widehat{\mathbb{E}} \big[f(\mathbf{W}_t) - f(e) \big] \right| \lesssim_f \sqrt{t},
$$

from which the result follows. $\qquad \square$

3.3 From Euclidean space to Riemannian manifolds

Let $(\mathcal{M}, g)$ be a Riemannian manifold of dimension d, without boundary. From Theorem A, the remainder of this section shows how to deduce the homogenisation property described in Theorem B, in the manifold setting. We emphasised in the introduction that anisotropic Brownian motion describes the random motion of a non-point-like object, with its own notion of local orientation. Such an object is represented by a point in the orthonormal frame bundle $O\mathcal{M}$ of $\mathcal{M}$, where its dynamics is described by a stochastic differential equation. We refer to Hsu's book [21] for a reference textbook on stochastic differential geometry.

In this subsection, we use Einstein summation convention: indices appearing twice are implicitly summed.

3.3.1 The orthonormal frame bundle

Denote by $z = (q, e)$ a generic point of the orthonormal frame bundle $O\mathcal{M}$ of $\mathcal{M}$, with $q \in \mathcal{M}$ and $e : \mathbb{R}^d \to T_q\mathcal{M}$, an orthonormal frame of $T_q\mathcal{M}$; we write $\pi : O\mathcal{M} \to \mathcal{M}$ for the canonical projection map. The Levi-Civita connection on $T\mathcal{M}$ induces a notion of horizontal vectors on $T\mathcal{M}$ or $O\mathcal{M}$. Let H stand for the horizontal lift operator, meaning the map $O\mathcal{M} \times \mathbb{R}^d \to TO\mathcal{M}$ uniquely characterised by the property that $\mathrm{H}_z(u) \in T_z O\mathcal{M}$ is horizontal and

$$
\mathrm{d}\pi_z \big(\mathrm{H}_z(u) \big) = e(u),
$$

for any $u \in \mathbb{R}^d$ and $z = (q, e) \in O\mathcal{M}$. Letting $(\epsilon_1, \ldots, \epsilon_d)$ be the canonical basis of $\mathbb{R}^d$, local coordinates q^i on $\mathcal{M}$ induce canonical coordinates on $O\mathcal{M}$ by writing

$$e_i := e(\epsilon_i) = e_i^j \frac{\partial}{\partial q^j}.$$

Denoting by Γ_{ij}^k the Christoffel symbols of the Levi-Civita connection associated with the above coordinates, the vector fields $\mathrm{H}(u)$ have the following expression.

$$\mathrm{H}_z(\epsilon_\alpha) = e_\alpha^i \frac{\partial}{\partial q^i} - \Gamma_{ij}^k(q) e_\alpha^i e_l^j \frac{\partial}{\partial e_l^k}.$$

3.3.2 Cartan's development map and anisotropic kinetic Brownian motion

Roughly speaking, Cartan's development map associates in its simplest form a $\mathcal{C}^1$ path in $\mathcal{M}$, started from $q_0 \in \mathcal{M}$, to any $\mathcal{C}^1$ path in the Euclidean space $\mathbb{R}^d$. Technically, given a $\mathcal{C}^1$ path $(x_t)_{t \geq 0}$ in $\mathbb{R}^d$, and $z_0 = (q_0, e_0) \in O\mathcal{M}$, the Cartan development of $(x_t)_{t \geq 0}$ on $\mathcal{M}$ is defined as the projection $(q_t)_{0 \leq t < T}$ on $\mathcal{M}$ of the horizontal $O\mathcal{M}$-valued path $(z_t) =: (q_t, e_t)_{0 \leq t < T}$ solution of the ordinary differential equation

$$\mathrm{d}z_t = \mathrm{H}_{z_t}(\mathrm{d}x_t), \quad \text{i.e.} \quad \dot{z}_t = \mathrm{H}_z(\dot{x}_t) \tag{3.3}$$

started from q_0, possibly up to some explosion time T. Note that the choice of $x : t \mapsto tu$ for some $u \in \mathbb{R}^d$ leads to q being a geodesic with initial condition $\dot{q}_0 = e_0(u)$; in particular, the development of X^σ tends to a geodesic with random initial condition as $\sigma \to 0$.

Classical stochastic analysis (in the Stratonovich sense) can be used to make sense of the preceding equation for x a semimartingale, defining Cartan's stochastic development — refer to Hsu's book [21] for a pedagogical account of the theory. For example, one of the many equivalent constructions of Brownian motion on $\mathcal{M}$ started at q_0 consists in developing a standard Euclidean Brownian motion. Accordingly, we define anisotropic Brownian motion on $\mathcal{M}$ as the development of the Euclidean Brownian motion with covariance $\mathrm{diag}(\gamma_1, \cdots, \gamma_d)$.

Anisotropic kinetic Brownian motion $(q_t^\sigma)_{0 \leq t < T}$ on $\mathcal{M}$ is the stochastic development of the anisotropic kinetic Brownian motion $(X_t^\sigma)_{t \geq 0}$ on $\mathbb{R}^d$; it is indexed by the speed parameter σ of its flat counterpart. This is a $\mathcal{C}^1$ random path which depends on the entire frame e_0 — its isotropic counterpart only depends *in law* on e_0, from symmetry properties of Wiener measure on $\mathbb{R}^d$. Although X^σ converges in law to an anisotropic Brownian motion B on $\mathbb{R}^d$, the poor regularity properties of the Itô solution map does not allow to conclude that anisotropic Brownian motion x^σ on $\mathcal{M}$ converges to projection on $\mathcal{M}$ of the solution of the equation

$$\mathrm{d}z_t = \mathrm{H}(z_t) \circ \mathrm{d}B_t.$$

This is exactly the kind of conclusion that rough paths theory provides.

3.3.3 Rough paths and rough differential equations with values in manifolds

We discuss a few results of rough paths theory with values in manifolds. These results are all classical, and their Euclidean counterparts can be found e.g. in [15] or [16]. Let $\mathcal{N}$ be a manifold, and, for a collection $A = (A_1, \cdots, A_n)$ of smooth vector fields on $\mathcal{N}$ and an initial condition $p \in \mathcal{N}$, consider the (deterministic) controlled differential equation

$$\mathrm{d}z_t = A(z_t)\mathrm{d}x_t, \quad z_0 = p$$

on $\mathcal{N}$, where x is a driving curve with values in $\mathbb{R}^n$. The equation makes sense whenever x is of class $\mathcal{C}^1$ (dividing each side by $\mathrm{d}t$, one might say), and if moreover x is of class $\mathcal{C}^2$, its solution is characterised by the fact that for any fixed $t \geq 0$ and $f : \mathcal{N} \to \mathbb{R}$ smooth with compact support,

$$f(z_t) = f(z_s) + (A_i f)(z_t)(x_t^i - x_s^i) + O(|t-s|^2)$$

as $s \to t$. Now if $\mathbf{X} = (X, \mathbb{X})$ is a rough path of Hölder regularity $1/3 < \gamma \leq 1/2$, we consider the following notion of solution: a continuous path $z : [0, T) \to \mathcal{N}$ is a solution of the rough differential equation

$$\mathrm{d}z_t = A(z_t)\mathbf{X}_{\mathrm{d}t}, \quad z_0 = p \tag{3.4}$$

if one can find some $a > 1$ such that any choice of $t \geq 0$ and $f : \mathcal{N} \to \mathbb{R}$ smooth with compact support yields

$$f(z_t) = f(z_s) + (A_i f)(z_t)(X_t^i - X_s^i) + (A_i A_j f)(z_t)\mathbb{X}_{ts}^{ij} + O(|t-s|^a)$$

as $s \to t$. This point of view is taken from [4, 5], in the mindset of [13]. We say that z explodes as $t \to T$ if z leaves any compact set.

In a probabilistic mindset, the fundamental remark is that, for $\mathbf{X}$ the Stratonovich rough path lift of some standard Brownian motion W, such a solution coincides almost surely with the solution of the Stratonovich equation

$$\mathrm{d}z_t = A(z_t) \circ \mathrm{d}W_t, \quad z_0 = p.$$

It is a striking feature of rough paths theory that not only does (3.4) admit a unique solution z for any (deterministic) rough path $\mathbf{X}$, in the above sense and up to some explosion time $T > 0$, but also the Itô-Lyons map $\mathbf{X} \mapsto z$ is continuous in the following sense. Fix d a Riemannian distance on $\mathcal{N}$. If $T' < T$ and $\varepsilon > 0$, there exists some $\delta > 0$ such that for any $\mathbf{X}'$ at rough path distance at most δ from $\mathbf{X}$, the solution z' of

$$\mathrm{d}z_t' = A(z_t')\mathbf{X}_{\mathrm{d}t}', \quad z_0' = p$$

is defined on $[0, T']$ and satisfies $d(z_t, z_t') < \varepsilon$ for all $0 \leq t \leq T'$.

This kind of continuity in enough to ensure convergence in law: namely, if $(\mathbf{X}^n)_{n \geq 0}$ is a family of random rough paths converging in law to $\mathbf{X}$ with respect to the rough path topology, then in a sense, the (random) solution z^n of (3.4) driven by $\mathbf{X}^n$ converges to the solution of that driven by $\mathbf{X}$. Let us make that point precise. Denote by $\widehat{\mathcal{N}}$ the one point compactification of $\mathcal{N}$ ($\widehat{\mathcal{N}} = \mathcal{N}$ if $\mathcal{N}$ is compact) and set C_p the space of continuous paths $z : [0, 1] \to \widehat{\mathcal{N}}$ starting at p such that $z_{t+} \equiv \infty$ whenever $z_t = \infty$. Fix d a Riemannian metric on $\mathcal{N}$ such that $d(p, p') \to \infty$ as $p' \to \infty$, and define on C_p the smallest topology containing, for any $\gamma \in C_p$ and $R, \varepsilon > 0$, the set of paths $z \in C_p$ satisfying

$$\max_{\substack{t \geq 0 \\ d(p, \gamma_t) \leq R}} d(z_t, \gamma_t) < \varepsilon.$$

The topology does not depend on d, and a sequence z^n of curves in C_p converges to z^∞ if and only if for all R, the curves $z_{\cdot \wedge \tau_R}^n$ stopped when they get at distance R of p converge uniformly to $z_{\cdot \wedge \tau_R}^\infty$. We can now state what one might call a theorem of continuity in law, in the following form.

Theorem 3.7. *For some fixed $1/3 < \gamma \leq 1/2$, let $(\mathbf{X}^n)_{n \geq 0}$ be a sequence of random γ-rough paths with values in $\mathbb{R}^d$, whose distributions converge weakly to that of $\mathbf{X}^\infty$. These processes might be defined on different probability spaces.*

Then, for any $0 \leq n \leq \infty$, there exists a unique random variable z^n with values in C_p such that it solves the rough differential equation

$$\mathrm{d}z_t^n = A(z_t^n)\mathbf{X}_{\mathrm{d}t}^n, \quad z_0^n = p$$

almost surely up to explosion, and the distributions of z^n converge to that of z^∞ with respect to the topology of C_p described above.

3.3.4 The interpolation result

The proof of Theorem B then follows from the rough path convergence of the rough path lift $\mathbf{X}^\sigma$ of anisotropic kinetic Brownian motion X^σ in $\mathbb{R}^d$, Theorem 3.6, and the continuity properties of the Itô-Lyons solution map to rough differential equations. As in [1], one needs to use the stochastic and geodesic completeness of $(\mathcal{M}, g)$ to conclude that the convergence of the $O\mathcal{M}$-valued development of anisotropic kinetic Brownian in $\mathbb{R}^d$ in not only local, but that weak convergence holds true; see Proposition 2.4.3 and Lemma 2.4.4 in [1]. Stochastic completeness refers here to the isotropic Brownian motion on $\mathcal{M}$. We implicitly use here the fact that for a complete and stochastically complete Riemannian manifold, the anisotropic Brownian motion on $\mathcal{M}$ is also stochastically complete.

4 A wider class of kinetic ergodic motions

In this section, we take a step back, and see what remains of Theorem B in a higher level of generality. Unfortunately, it is difficult to state a general theorem on kinetic ergodic motions without hiding the ideas under a layer of abstract machinery, unnecessary in most applications. Instead, we chose in this last section to introduce basic notations and state a prototypical result first, which is in essence a reformulation of Theorem C in the introduction, and then to discuss its hypotheses through a variety of examples, which we believe convey the versatility of the method.

4.1 From kinetic Brownian motion to kinetic ergodic processes

We introduce a few notations, which will be useful in the following section. Suppose that $(v_t^\sigma)_{t \geq 0}$ is of the form $v_t^\sigma = I(\overline{v}_{\sigma^2 t})$, with $(\overline{v}_t)_{t \geq 0}$ a càdlàg Markov process with values in some manifold $\mathcal{W}$ and $I : \mathcal{W} \to \mathbb{R}^d$ bounded continuous — for instance, Theorem B deals with the case $I : \mathcal{W} = \mathbb{S}^{d-1} \hookrightarrow \mathbb{R}^d$ and $\overline{v}$ the anisotropic spherical Brownian motion with time scale 1. Because the path $(x_t^\sigma)_{t \geq 0}$ integrating the velocity is Lipschitz, its development on a Riemannian manifold is well-defined, and the objects described in Theorem C make sense. Note also that, because the velocity is locally bounded, the rough path lift used in the proof is well-defined as well; see Remark 3.4.

This subsection is devoted to the proof of the following rewriting of Theorem C. We discuss its hypotheses in the following Subsection 4.2.

Theorem 4.1 (Main Theorem C)**.** *Let $(\mathcal{M}, g)$ be a Riemannian manifold of dimension d, and $(q_t^\sigma)_{t \geq 0}$ a process on $\mathcal{M}$ whose velocity $\dot{q}_t^\sigma \in T_{q_t}\mathcal{M}$ has image $v_t^\sigma \in T_{q_0}\mathcal{M} \simeq \mathbb{R}^d$ under the inverse stochastic parallel transport along q. Suppose that, for some càdlàg Markov process $\overline{v}$ on a manifold $\mathcal{W}$, $(v_t^\sigma)_{t \geq 0}$ is the continuous image of $(\overline{v}_{\sigma^2 t})_{t \geq 0}$, i.e. $v_t^\sigma = I(\overline{v}_{\sigma^2 t})$ with $I : \mathcal{W} \to T_{q_0}\mathcal{M}$ bounded continuous. Suppose that $\overline{v}$ admits an invariant measure μ such that under $\mathbb{P} = \mathbb{P}_\mu$,*

1. *equation (1.4) holds with $\mathcal{F}_{[a,b]}$ the σ-algebra generated by $\{I(\overline{v}_t)\}_{a \leq t < b}$;*

2. *for all $1 \leq i \leq d$, the reflections $(v^1, \cdots, v^{i-1}, -v^i, v^{i+1}, \cdots, v^d)$ have the same distribution as $v = v^\sigma = (v^1, \cdots, v^d)$ for some, hence all, $\sigma > 0$.*

Then as $\sigma \to \infty$, the time rescaled process $(q^\sigma_{\sigma^2 t})_{t \in [0,1]}$ converges in law to an anisotropic Brownian motion on $\mathcal{M}$ with covariance $\mathrm{diag}(\gamma_1, \cdots, \gamma_d)$,

$$\gamma_i := \int_0^\infty \mathbb{E}\left[I(\overline{v}_0)^i I(\overline{v}_t)^i \right] \mathrm{d}t.$$

Remark 4.2. Throughout our study, we have worked at equilibrium, with $\mathbb{P} = \mathbb{P}_\mu$. Although it simplifies the proofs, it is merely a cosmetic concern in the case of kinetic Brownian motion. In fact, under the assumption (2.2), Theorem 4.1 holds for any $\mathbb{P}_\lambda$: see Proposition 4.3 below. For instance, it will be the case in examples 4.2.2 and, to some extent, 4.2.6 below. It is not clear whether the result should hold without this additional property.

To establish Theorem 4.1, let us review the ingredients of the proof of Theorem B. The tightness results, more specifically Corollary 3.3, are essentially a consequence of Lemma 2.6. It holds whenever (1.4) is satisfied (condition (1)), $I(\overline{v}_0)$ is centred (condition (2)) and I is bounded. On the other hand, the convergence towards Brownian motion relies, in addition, on the symmetry property (condition (2)) and independence of the increments. Equation (1.4) ensures the latter, so that the proof of Theorem 4.1 is essentially that of Theorem B.

Proposition 4.3. *Replace condition* (1) *in Theorem 4.1 by the following variant of* (2.2). *There exists some mixing time $\tau > 0$ such that for all $x \in \mathcal{W}$ and $t > 0$,*

$$\| P_t^* \delta_x - \mu \|_{\mathrm{TV}} \le f(x) \exp(-t/\tau) \tag{4.1}$$

for some function $f : \mathcal{W} \to \mathbb{R}_+$ integrable with respect to μ. Let the initial distribution λ be any probability measure on $\mathcal{W}$ such that $\int f \mathrm{d}\lambda < \infty$.

Then under the measure $\mathbb{P}_\lambda$, for $(q_t^\sigma)_{t \ge 0}$ defined as in Theorem 4.1 above and as $\sigma \to \infty$, the time rescaled process $(q^\sigma_{\sigma^2 t})_{t \in [0,1]}$ also converges in law to an anisotropic Brownian motion on $\mathcal{M}$ with the same covariance $\mathrm{diag}(\gamma_1, \cdots, \gamma_d)$.

Proof. It is enough to show the convergence of the Euclidean rough paths $(\mathbf{X}^\sigma)_{\sigma > 0}$.

Tightness We claim that Proposition 2.4 holds for $\mathbb{E}_\lambda$. Indeed, by the same arguments, we see that

$$|\mathbb{E}_{P_u^* \delta_x}[G] - \mathbb{E}_\mu[G]| \le |G|_\infty f(x) \, \mathrm{e}^{-u/\tau} \tag{4.2}$$

holds in lieu of (2.3). From this we deduce

$$|\mathbb{E}_{P_{t-s}^* \delta_x}[G] - \mathbb{E}_{P_t^* \lambda}[G]| \le |\mathbb{E}_{P_{t-s}^* \delta_x}[G] - \mathbb{E}_\mu[G]| + |\mathbb{E}_\mu[G] - \mathbb{E}_{P_t^* \lambda}[G]|$$

$$\le \left(f(x) + \left(\int f \mathrm{d}\lambda \right) \cdot \mathrm{e}^{-s/\tau} \right) |G|_\infty \, \mathrm{e}^{-(t-s)/\tau},$$

which is enough for rest of the proof to hold. It is then an easy exercise to adapt the proof of Corollary 2.5, and from this point every idea leading to tightness is the same, even if some care must be given to non-stationarity in the actual computations, e.g. regarding equation (3.1).

Brownian limit Let $\widehat{\mathbb{P}}_\mu$ be the law of the Brownian rough path on $\mathrm{RP}(\gamma)$, and $\widehat{\mathbb{P}}_\lambda$ a limit point of the laws of $\mathbf{X}^\sigma$ under $\mathbb{P}_\lambda$. We only need to show that $\widehat{\mathbb{P}}_\lambda = \widehat{\mathbb{P}}_\mu$.

Define the translation operator T_h on $\mathrm{RP}(\gamma)$ as

$$T_h(Y, \mathbb{Y}) := (Y_{h+\cdot} - Y_h, \mathbb{Y}_{h+\cdot, h+\cdot}).$$

Now, for any continuous bounded map $F : \mathcal{C}([0,1], G) \to \mathbb{R}$ and $\varepsilon > 0$, equation (4.2) gives

$$|\mathbb{E}_\lambda[F(T_\varepsilon \mathbf{X}^\sigma)] - \mathbb{E}_\mu[F(T_\varepsilon \mathbf{X}^\sigma)]| \leq \lambda(f)|F|_\infty \, \mathrm{e}^{-\sigma^4 \varepsilon / \tau},$$

which, taking limits along a proper subsequence, implies that

$$\widehat{\mathbb{E}}_\lambda[F(T_\varepsilon \mathbf{X})] = \widehat{\mathbb{E}}_\mu[F(T_\varepsilon \mathbf{X})].$$

But $T_\varepsilon \mathbf{X} \to \mathbf{X}$ in $\mathcal{C}^0([0,1], G)$ for fixed $\mathbf{X}$, so the above equation holds for $\varepsilon = 0$, and $\widehat{\mathbb{P}}_\lambda$ is the law of the announced anisotropic Brownian motion. Note that $T_\varepsilon \mathbf{X}$ has no reason to converge to $\mathbf{X}$ in the rough path topology, so tightness had to be proved beforehand. $\quad \square$

4.2 Discussion of the hypotheses

Because we introduced Theorem 4.1 merely as a reformulation of Theorem B with highlight on the key assumptions, there is much room for improvement. Below are a few examples where the hypotheses can be weakened; in theory, none of these methods are mutually exlusive, although the manipulations might get tedious when applied to actual data.

4.2.1 Asymmetrical velocity — spinning motion

We show in this first example illustrates that the symmetry condition (2) in Theorem 4.1 above is indeed necessary. Assuming only that $I(\overline{v})$ is centred, the existence and Brownian behaviour of the Euclidean limit follows from the tightness result stated in Corollary 3.3 and the characterisation of Proposition 3.5, whose proofs hold with no change whatsoever; however, the limit rough path need not be Brownian, as is the case in what follows.

Set $I : \mathcal{W} = \mathbb{R}/2\pi\mathbb{Z} \to \mathbb{C} \simeq \mathbb{R}^2$ the exponential $v \mapsto \mathrm{e}^{iv}$, and define $\overline{v}$ as the spinning motion

$$\mathrm{d}\overline{v}_t = \mathrm{d}t + \mathrm{d}W_t, \quad \text{i.e.} \quad \overline{v}_t = \overline{v}_0 + t + W_t \pmod{2\pi},$$

where W is a standard Brownian motion on $\mathbb{R}$. Its dynamics is of course very simple: it admits a unique invariant measure $\mu(\mathrm{d}v) = \frac{1}{2\pi}\mathrm{d}v$ and satisfies equation (2.2), so all hypotheses but condition (2) in Theorem 4.1 above are satisfied.

As described above, the laws of $(X^\sigma)_{\sigma>0}$ do converge to that of a Brownian process. As of those of the lifts $(\mathbf{X}^\sigma)_{\sigma>0}$, however, some drift appears in the limit. Indeed, setting $\mathbb{A}^\sigma$ the antisymmetric part of $\mathbb{X}^\sigma$,

$$(\mathbb{A}_{t0}^\sigma)^{12} = \frac{1}{2\sigma^4} \int_0^{\sigma^4 t} \int_0^s \sin(\overline{v}_s - \overline{v}_u)\mathrm{d}s\mathrm{d}u = \int_0^\infty \int_0^\infty \frac{1}{2\sigma^4} \mathbf{1}_{u+\tau \leq \sigma^4 t} \sin(\overline{v}_{u+\tau} - \overline{v}_u)\mathrm{d}u\mathrm{d}\tau,$$

so we get

$$\mathbb{E}\big[(\mathbb{A}_{t0}^\sigma)^{12}\big] = \int_0^\infty \int_0^\infty \frac{1}{2\sigma^4} \mathbf{1}_{u+\tau \leq \sigma^4 t} \sin(\tau)\, \mathrm{e}^{-\tau/2}\mathrm{d}u\mathrm{d}\tau = \frac{1}{2} \int_0^\infty \left(t - \frac{\tau}{\sigma^4}\right)_+ \sin(\tau)\, \mathrm{e}^{-\tau/2}\mathrm{d}\tau$$

with $(\cdot)_+$ the positive part. The limit is a non-zero linear function of t, so the limit of the lifts cannot be Brownian.

In particular, the homogenisation result does not hold for its manifold-valued counterpart, and one should not consider this example as an isolated pathological case. In the common 'rolling without slipping' analogy used to described stochastic development of Brownian motion, one might think of the resulting non-Brownian effect as a force rotating the paper around the contact point, so that the path on the manifold may have a tendency to lean to one side.

Such drift phenomena in the Lévy area have arisen and been studied in different works recently, particularly in the context of random walks. See e.g. the articles [31, 30] of Lopusanschi and Simon, and those of Ishiwata, Kawabi and Namba, [22, 23]. In the continuous setting, it has been described by Friz, Gassiat and Lyons in [14] as the small-mass limit of the behaviour of a charged particle in a magnetic field.

4.2.2 Velocity with jumps — random flight

The so-called random flight is perhaps the most elementary situation where the velocity has jumps. Studied by Pinsky under the name of isotropic transport process in [33], it is a piecewise-geodesic motion on a Riemannian manifold M where the direction is sampled uniformly on the unit tangent sphere, and updated at exponential times. It is shown in [33] to exhibit a limiting Brownian behaviour under suitable normalisation. In terms of our I and $\overline{v}$, we describe it by setting $I : \mathcal{W} = \mathbb{S}^{d-1} \hookrightarrow \mathbb{R}^d$ and $\overline{v}$ a pure jump process, with rate 1 and uniform measure. In this case, the mixing property (1.4) is a consequence of the stronger statement (2.2) that the dynamics converges exponentially fast to equilibrium in total variation, in the same way we treated anisotropic Brownian motion, so that our Theorem 4.1 applies readily and the limit is indeed Brownian. There are no complications in dealing with jumps.

Because the velocity process is isotropic, the limit covariance $\mathrm{diag}(\gamma_1, \cdots, \gamma_d)$ is proportional to Id. Setting T the first jump time,

$$\gamma_i = \frac{2}{d} \int_0^\infty \mathbb{E}[\overline{v}_0 \cdot \overline{v}_t] \mathrm{d}t = \frac{2}{d} \int_0^\infty \mathbb{P}(T \leq t) \mathrm{d}t = \frac{2}{d},$$

and we recover the result of [33].

4.2.3 Discrete time — Donsker invariance principle for random walks

In certain situations, the ergodic properties of the system are most easily described in deterministic discrete time. An example, studied in [8] by Breuillard, Friz and Huesmann, is that of random walks. If $(Y_k)_{k \geq 0}$ is a sequence of independent *bounded* random variables with values in $\mathbb{R}^d$, *symmetric* in the sense that their common law is invariant with respect to the reflections as described in condition (2) in Theorem 4.1 above, we consider the piecewise linear process W that, on each interval $[k, k+1]$, is affine and increases by Y_k. In other words, we express it as

$$W : t \mapsto \sum_{k < \lfloor t \rfloor} Y_k + (t - \lfloor t \rfloor) Y_n.$$

The classical invariance principle of Donsker states that the rescaled processes $W^\sigma : t \mapsto W_{\sigma^4 t}/\sigma^2$ converge in law to an anisotropic Brownian motion, with respect to the uniform convergence on compact sets (in most expositions, the convergence is proved for $\sigma^4 = N$, as $N \to \infty$). The result of [8] strengthens it to rough path convergence.

Let us translate this dynamics in our framework. Set $\mathcal{W} = \mathbb{R}/\mathbb{Z} \times \mathbb{R}^d$, $I : (\alpha, y) \mapsto y$, and define the dynamics of $(\overline{v}_t)_{t \geq 0} = (\alpha_t, y_t)_{t \geq 0}$ as follows. Given initial conditions $(\alpha_0, y_0) \in [0, 1) \times \mathbb{R}^d$, α grows continuously with rate 1, i.e. $\alpha_t = \alpha_0 + t \pmod 1$, whereas y stays constant on time intervals of length 1, then jumps independently of the past according to the law of Y_k, i.e. $y_t = Y_{\lfloor t - \alpha_0 \rfloor}$ with the convention $Y_{-1} = y_0$. Under the initial condition $\delta_0 \otimes \mathcal{L}(Y_0)$, we see that the law of X^σ is exactly that of W^σ.

The process $\overline{v}$ is Markovian, although not Feller, and admits an invariant measure $\mathrm{Unif}(\mathbb{R}/\mathbb{Z}) \otimes \mathcal{L}(Y_0)$. Because it is not ergodic, there is no hope for equation (2.2) to hold. Maybe surprisingly, even if I kills the non-mixing coordinate, it is also false that

condition (1) of Theorem 4.1 holds: in the case where Y has no atoms, take P to be the first jump time in $[0, 1]$, and F the first jump time in $[n, n+1]$. However, it is true for any $\mu_\alpha := \delta_\alpha \otimes \mathcal{L}(Y_0)$, with constants independent of α: indeed, it is obvious that for $t > 1$ and any probability law λ on $\mathbb{R}^d$,

$$P_t^*(\delta_\alpha \otimes \lambda) = \mu_{\alpha+t} = P_t^* \mu_\alpha$$

holds in lieu of (2.2). Remarkably, nothing more than this is needed throughout the proof. It should be clear that Proposition 2.5 holds for any μ_α, and that tightness follows in the same fashion. Independence of increments, as stated in Propositions 3.5 and Theorem 3.6, hides no difficulty either. It is true that one has to be careful about the limit variance in Proposition 3.5, because the Markov property is used in a crucial way. In our case, for any $\alpha \in [0, 1)$ and $\sigma > 1$, we end up with

$$
\begin{aligned}
\mathbb{E}_{\mu_\alpha}\big[((X_1^\sigma)^i)^2\big] &= \frac{1}{\sigma^4} \int_0^{\sigma^4} \int_0^{\sigma^4} \mathbb{E}_{\mu_\alpha}[y_s^i y_t^i] \, \mathrm{d}s \, \mathrm{d}t \\
&= \sum_{n \geq -1} \frac{1}{\sigma^4} \int_0^{\sigma^4} \int_0^{\sigma^4} \mathbf{1}_{n+\alpha \leq s, t < n+1+\alpha} \mathbb{E}_{\mu_\alpha}[y_s^i y_t^i] \, \mathrm{d}s \, \mathrm{d}t \\
&= \mathbb{E}[|Y_0^i|^2] \cdot \frac{(1-\alpha)^2 + \lfloor \sigma^4 + \alpha - 1 \rfloor + \{\sigma^4 + \alpha - 1\}^2}{\sigma^4}
\end{aligned}
$$

with $\{\cdot\}$ the fractional part. In the limit, the variance converges to $\mathbb{E}[|Y_0^i|^2]$, so the result of Theorem B, and in particular the rough path strengthening of the Donsker invariance principle, holds with covariance $\mathbb{E}[Y_0 Y_0^*]$, in accordance with [8].

Surprisingly enough, the symmetry condition (2) is not mandatory here: see [8]. Intuitively, the increments would have to work together to spin in some privileged direction, but this is prevented by independence. In particular, a drift in the antisymmetric part of $\mathbb{X}^\sigma$ that does not vanish in the limit, as mentioned in example 4.2.1 above, must come from additional structure: in [31], the hidden Markov chain; in [30], the underlying directed graph; etc.

Note that in the case of random walks, as a consequence of the work of Chevyrev, see [9, Example 5.8], convergence of X^σ as stated in Proposition 3.5 in enough to ensure convergence of $\mathbf{X}^\sigma$ to some random rough path. It is not clear from this approach, however, that this limit is indeed Brownian.

4.2.4 Discrete time and correlation — Donsker invariance principle for Markov chains

The reader may have noticed that in the above example 4.2.3, independence of the variables $(Y_k)_{k \geq 0}$ is a bit much, and one could work with covariances vanishing exponentially fast. Suppose for instance that $(Y_k)_{k \geq 0}$ is a time-homogeneous Markov chain with invariant measure μ with compact support, whose correlations decrease as $\mathrm{e}^{-k/\tau}$, $\tau > 0$; namely, letting Q be the transition kernel of Y,

$$\|\delta_y Q^k - \mu\|_{\mathrm{TV}} \lesssim \mathrm{e}^{-k/\tau}$$

for all y in the support of μ. Then, setting $\mu_0 := \delta_0 \otimes \mu$, we get, for any probability measure λ with $\mathrm{Supp}\, \lambda \subset \mathrm{Supp}\, \mu$,

$$\|P_t^*(\delta_0 \otimes \lambda) - P_t^* \mu_\alpha\|_{\mathrm{TV}} = \left\| \int (\delta_y Q^{\lfloor t \rfloor} - \mu)\lambda(\mathrm{d}y) \right\|_{\mathrm{TV}} \lesssim \mathrm{e}^{-t/\tau}.$$

Again, this inequality can be substituted for equation (2.2) in the proof of Proposition 2.4, and under the same symmetry condition as above, the convergence result still holds true.

Examples of such Markov chains are any aperiodic irreducible finite state Markov chain; or any Markov chain with transition kernel $Q(y, \mathrm{d}y')$ absolutely continuous with respect to some measure ν, and such that $\frac{\mathrm{d}Q(y,\cdot)}{\mathrm{d}\nu}$ is bounded below by a positive constant $m > 0$, uniformly in y, y'. Note however that the symmetry condition (2) of Theorem 4.1 is a bit stronger than in the independent case, since we need the reflections to leave the law of the whole sequence invariant.

4.2.5 Time-dependent randomness — time-dependent Brownian motion

The way we wrote our convergence theorems is ill-suited to treat time-dependent randomness, i.e. when the motion of the velocity is not homogeneous in time. However, there are cases where randomness can be somewhat dissociated from the time dependence, and our methods do in fact yield interesting convergence results. In the present example, we set to recover, in the limit, the Brownian motion on a manifold $\mathcal{M}$ endowed with a time-dependent metric g_t, as introduced in [3] by Arnaudon, Coulibaly and Thalmaier.

Such an approach has already been set up in [25], in a similar fashion as the random flight described in example 4.2.2 above. The idea is to freeze the metric in small time intervals $[t_i, t_{i+1}]$, say of size $1/\sigma^4$, over which the movement q is purely geodesic with respect to the metric g_{t_i}, the initial condition being chosen uniformly at t_i on the unit g_{t_i}-sphere of the tangent space of $\mathcal{M}$ at q_{t_i}. Suitably renormalised, this process converges to the time-dependent Brownian motion described above. We introduce a similar random flight which lets the metric vary continuously, and may be considered more natural in this respect, then prove its convergence to time-dependent Brownian motion.

We begin by describing time-dependent Brownian motion and its surroundings. Suppose g_t is smooth, as a function on $\mathbb{R}_+ \times T\mathcal{M} \otimes T\mathcal{M}$. Let $F\mathcal{M}$ be the frame bundle over $\mathcal{M}$, and choose a point $q_0 \in \mathcal{M}$ together with a g_0-orthonormal frame e_0 of $T_{g_0}\mathcal{M}$. For a $\mathcal{C}^1$ path $(x_t)_{t \geq 0}$ in $\mathbb{R}^d$, we define the time-dependent development of x as the solution $(z_t)_{t \geq 0} = (q_t, e_t)_{t \geq 0}$ of the following equation, whose terms we describe below.

$$\mathrm{d}z_t = \mathrm{H}_{t,z_t}(\mathrm{d}x_t) - \frac{1}{2}\frac{\partial g_t}{\partial t}(u_t \epsilon_i, u_t \epsilon_j)\mathrm{V}_{z_t}^{ij}\mathrm{d}t, \quad z_0 = (q_0, e_0). \tag{4.3}$$

We use Einstein notation. As in Section 3.3.1, $(\epsilon_1, \cdots, \epsilon)$ is the canonical basis of $\mathbb{R}^d$, and the $\mathrm{H}_{t,z}\epsilon_i$, resp. V_z^{ij}, are the canonical horizontal vector fields, resp. vertical vector fields. Note that because the metric g is time-dependent, the associated horizontal vector fields H must depend on t as well. In coordinates,

$$\mathrm{H}_{t,z}(\epsilon_\alpha) = e_\alpha^i \frac{\partial}{\partial q^i} - \left(\Gamma_t(q)\right)_{ij}^k e_\alpha^i e_l^j \frac{\partial}{\partial e_l^k}, \qquad \mathrm{V}_z^{ij} = e_j^k \frac{\partial}{\partial e_i^k}.$$

If we compare (4.3) to (3.3), the added vertical fields are there to ensure that e_t is at all times orthonormal for g_t. We refer to [12] for an insight about why this definition is a sensible choice.

In particular, the time-dependent geodesics are the solutions of the equation associated to $x_t = tu$ for some fixed $u \in \mathbb{R}^d$, and the time-dependent Brownian motion is the solution driven by some standard Brownian motion W in the Stratonovich sense, or, equivalently, by the standard Stratonovich rough path $\mathbf{W}$ in the rough sense.

Note that we did not discuss time-dependent rough differential equations in Section 3.3.3. In the case of an equation driven by a $\mathcal{C}^1$ control x, the standard technique is of course to consider $t \mapsto (t, x_t)$ as the control. The same trick works with rough paths:

associated to any rough path $\mathbf{Y} = (Y, \mathbb{Y})$ is a canonical lift $\widehat{\mathbf{Y}}$ of $t \mapsto (t, Y_t)$ compatible with $\mathbf{Y}$. The solution of time-dependent rough differential equations is then well-defined. In what follows, we will also use the fact that $\mathbf{Y} \mapsto \widehat{\mathbf{Y}}$ is continuous in the rough path topology, so $\widehat{\mathbf{X}}^\sigma \to \widehat{\mathbf{X}}$ in law whenever $\mathbf{X}^\sigma \to \mathbf{X}$ in law.

We define a kind of interpolated random walk on $\mathcal{M}$ whose limit will be the Brownian motion described above. Fix $\sigma > 0$, and define W^σ successively on each interval $[s, t] = [\frac{n}{\sigma^4}, \frac{n+1}{\sigma^4}]$ as follows: ξ_n^σ is chosen independently of all the rest according to the uniform measure on the unit g_s-sphere of $T_{W_s^\sigma} \mathcal{M}$, and W^σ on $[s, t]$ is a time-dependent geodesic in the above sense, with initial condition $\dot{W}_s^\sigma = \sqrt{d}\, \xi_n^\sigma$.

As in the previous example, there is a direct equivalent of this dynamics in our framework. Set $\mathcal{W} = \mathbb{R}/\mathbb{Z} \times \mathbb{S}^{d-1}$ and $I : (\alpha, y) \hookrightarrow \sqrt{d} y$, following the same dynamics as in 4.2.3, with Y_0 uniformly distributed on $\mathbb{S}^{d-1}$. We choose the initial condition to be $\delta_0 \otimes \mathrm{Unif}(\mathbb{S}^{d-1})$; for the same reasons as in example 4.2.3 above, $(\mathbf{X}^\sigma)_{\sigma > 0}$ converges to the Brownian rough path $\mathbf{X}$ with covariance $d\, \mathbb{E}[Y_0 Y_0^*] = \mathrm{Id}$.

Everything described so far is essentially time-invariant — the time-dependence appears when we use this family of rough paths to describe a motion on $\mathcal{M}$. Fix $q_0 \in \mathcal{M}$, and e_0 a g_0-orthonormal frame of $T_{q_0}\mathcal{M}$. Define the solution $(z_t)_{t \geq 0} = (q_t, e_t)_{t \geq 0}$ (up to explosion) on the frame bundle $F\mathcal{M}$ of equation (4.3) driven by $\mathbf{X}$, in the rough sense.

By definition, q_t defined as above is the Brownian motion associated to the time-dependent metric g_t, as described in [3]. If we set $z^\sigma = (q^\sigma, u^\sigma)$ the solution of the equation driven by $\mathbf{X}^\sigma$, we get instead $q^\sigma = W^\sigma$ in law. The convergence of $\mathbf{X}^\sigma$, together with the general theory of rough paths (see Theorem 3.7), ensures that q^σ, hence W^σ, converges in law to the time-dependent Brownian motion q.

4.2.6 Velocity with unbounded support — Langevin process

We conclude with an example where the velocity v has unbounded support, which goes against the implicit assumption that I be bounded. We consider the process with anisotropic Ornstein-Uhlenbeck velocity, i.e. satisfying

$$\mathrm{d}\overline{v}_t = -\overline{v}_t \mathrm{d}t + \mathrm{d}B_t$$

for B an anisotropic Brownian motion of covariance Σ. In the isotropic case, it is a simple scalar example of the hypoelliptic Laplacian of Bismut, and convergence towards Brownian motion can be seen as a geometric analogue of the interpolation described by the latter between the Laplace operator and the geodesic generator; see [7]. The anisotropic convergence is also treated in [6].

When the parameter σ is introduced, we get a velocity v_t satisfying

$$\mathrm{d}v_t = -\sigma^2 v_t \mathrm{d}t + \sigma \mathrm{d}B_t,$$

and we want to prove that the associated position process converges to Brownian motion, as $\sigma \to \infty$ and under suitable time change. Here, $I : \mathcal{W} = \mathbb{R}^d \to \mathbb{R}^d$ is simply the identity, and hence does not quite fit the boundedness hypothesis of Theorem C. However, it is well known that $\overline{v}$ admits as an invariant measure the Gaussian distribution $\mu = \mathcal{N}(0, \frac{1}{2}\Sigma)$ with covariance $\frac{1}{2}\Sigma$. Using the coupling $B_t' = -B_t$, it is known, and not difficult to see, that

$$\|P_t^* \delta_x - \mu\|_{\mathrm{TV}} \lesssim (1 \vee |x|)\, \mathrm{e}^{-t},$$

from whence, because $1 \vee |x|$ is in $L^1(\mu)$, we derive Proposition 2.4; see Proposition 4.3.

In our proof, boundedness of the velocity is essentially used twice: for proving the decorrelation of coordinates in Proposition 2.5, and to show that the variance of the limit must be the limit of the variances in 3.5. Because μ has moments of all order, the latter

will add no difficulty — in fact, any moment of order > 2 would suffice. As for the former, it is a bit trickier. We use the following variation of Proposition 2.5.

Proposition 4.4. *Fix some $\varepsilon > 0$ and some positive integer $n \in \mathbb{N}^*$. There exists $\tau' = \tau'(\tau, n, \varepsilon) > 0$ such that under $\mathbb{P} = \mathbb{P}_\mu$, and for any indices $1 \leq j_1, \cdots, j_n \leq d$ and times $s_1, \cdots, s_n \geq 0$,*

$$\left| \mathbb{E}\left[v_{s_1}^{j_1} \cdots v_{s_1 + \cdots + s_n}^{j_n} \right] \right| \lesssim |v_0^{j_1}|_{L^{n+\varepsilon}} \cdots |v_0^{j_n}|_{L^{n+\varepsilon}} \, \mathbf{e}^{-\Delta/\tau'}.$$

We give only hints of the proof. In the spirit of the proof of Proposition 2.5, set

$$V_- := \prod_{1 \leq k < k_0} \left(v_{t_k}^{j_k} / |v_0^{j_k}|_{L^{n+\varepsilon}} \right)$$

and similarly for V_0 and V_+. Write $V_* = W_* + R_*$ with $W_* := V_* \mathbf{1}_{|V_*| \geq M}$; for $M = \exp(\eta \Delta)$ with $\eta > 0$ small enough, the proof of Proposition 2.5 applied to W_*, together with a careful handling of the remainder R_*, are enough to get to the above result. It automatically implies Lemma 2.6, since μ has moments of all order, hence the conclusion of Theorem B.

Note that the treatment of unboundedness is not specifically designed for the Langevin process, so it can be applied to the above study of the random walk as well. Moreover, it is not necessary for all moments to exist: moments of order $\alpha > 2/(1 - 2\gamma)$ are enough to ensure tightness in $\mathsf{RP}(\gamma)$; this means $\alpha > 6$ if one only cares about the convergence of the development with respect to the uniform topology. Indeed, our proof, enhanced by the above corollary, will hold with moments of order $2n > 2/(1 - 2\gamma)$ for any positive integer n; but adding an easy truncation argument at the beginning of the proofs of Lemma 3.1 and 3.2 will strengthen the result to non even integral moments. In this respect, our moment assumption is a bit weaker than that of [8], when we stay in the symmetrical case.

References

[1] J. Angst, I. Bailleul, and C. Tardif, *Kinetic Brownian motion on Riemannian manifolds*, Electron. J. Probab. **20** (2015), no. 110, 40. MR-3418542

[2] D. Applebaum, *Probability on compact Lie groups*, Probability Theory and Stochastic Modelling, vol. 70, Springer, Cham, 2014, With a foreword by Herbert Heyer. MR-3243650

[3] M. Arnaudon, K. A. Coulibaly, and A. Thalmaier, *Brownian motion with respect to a metric depending on time: definition, existence and applications to Ricci flow*, C. R. Math. Acad. Sci. Paris **346** (2008), no. 13-14, 773–778. MR-2427080

[4] I. Bailleul, *A stochastic approach to relativistic diffusions*, Ann. Inst. Henri Poincaré Probab. Stat. **46** (2010), no. 3, 760–795. MR-2682266

[5] I. Bailleul, *Flows driven by rough paths*, Rev. Mat. Iberoam. **31** (2015), no. 3, 901–934. MR-3420480

[6] J. Birrell, S. Hottovy, G. Volpe, and J. Wehr, *Small mass limit of a Langevin equation on a manifold*, Ann. Henri Poincaré **18** (2017), no. 2, 707–755. MR-3596775

[7] J.-M. Bismut, *Hypoelliptic Laplacian and probability*, J. Math. Soc. Japan **67** (2015), no. 4, 1317–1357. MR-3417500

[8] E. Breuillard, P. Friz, and M. Huesmann, *From random walks to rough paths*, Proc. Amer. Math. Soc. **137** (2009), no. 10, 3487–3496. MR-2515418

[9] I. Chevyrev, *Random walks and Lévy processes as rough paths*, Probab. Theory Related Fields **170** (2018), no. 3-4, 891–932. MR-3773803

[10] I. Chevyrev, P. K. Friz, A. Korepanov, I. Melbourne, and H. Zhang, *Multiscale systems, homogenization, and rough paths*, Probability and analysis in interacting physical systems, Springer Proc. Math. Stat., vol. 283, Springer, Cham, 2019, pp. 17–48. MR-3968507

[11] M. Christensen and J. B. Pedersen, *Diffusion in inhomogeneous and anisotropic media*, The Journal of Chemical Physics **119** (2003), 5171–5175.

[12] K. A. Coulibaly-Pasquier, *Brownian motion with respect to time-changing Riemannian metrics, applications to Ricci flow*, Ann. Inst. Henri Poincaré Probab. Stat. **47** (2011), no. 2, 515–538. MR-2814421

[13] A. M. Davie, *Differential equations driven by rough paths: An approach via discrete approximation.*, AMRX, Appl. Math. Res. Express **2007** (2008), 40 (English). MR-2387018

[14] P. Friz, P. Gassiat, and T. Lyons, *Physical Brownian motion in a magnetic field as a rough path*, Trans. Amer. Math. Soc. **367** (2015), no. 11, 7939–7955. MR-3391905

[15] P. K. Friz and M. Hairer, *A course on rough paths*, Universitext, Springer, Cham, 2014, With an introduction to regularity structures. MR-3289027

[16] P. K. Friz and N. B. Victoir, *Multidimensional stochastic processes as rough paths*, Cambridge Studies in Advanced Mathematics, vol. 120, Cambridge University Press, Cambridge, 2010, Theory and applications. MR-2604669

[17] L. Gross, *Abstract Wiener spaces*, Proc. Fifth Berkeley Sympos. Math. Statist. and Probability (Berkeley, Calif., 1965/66), Vol. II: Contributions to Probability Theory, Part 1, Univ. California Press, Berkeley, Calif., 1967, pp. 31–42. MR-0212152

[18] L. Gross, *Abstract Wiener measure and infinite dimensional potential theory*, Lectures in Modern Analysis and Applications, II, Lecture Notes in Mathematics, Vol. 140. Springer, Berlin, 1970, pp. 84–116. MR-0265548

[19] M. Heidernätsch, M. S. Bauer, and G. Radons, *Characterizing n-dimensional anisotropic brownian motion by the distribution of diffusivities.*, The Journal of Chemical Physics **139** (2013), no. 18, 184105.

[20] D. P. Herzog, S. Hottovy, and G. Volpe, *The small-mass limit for Langevin dynamics with unbounded coefficients and positive friction*, J. Stat. Phys. **163** (2016), no. 3, 659–673. MR-3483250

[21] E. P. Hsu, *Stochastic analysis on manifolds*, Graduate Studies in Mathematics, vol. 38, American Mathematical Society, Providence, RI, 2002. MR-1882015

[22] S. Ishiwata, H. Kawabi, and R. Namba, *Central limit theorems for non-symmetric random walks on nilpotent covering graphs: Part i*, 2018.

[23] S. Ishiwata, H. Kawabi, and R. Namba, *Central limit theorems for non-symmetric random walks on nilpotent covering graphs: Part ii*, 2018.

[24] N. van Kampen, *Diffusion in inhomogeneous media*, Journal of Physics and Chemistry of Solids **49** (1988), 673–677. MR-0909885

[25] K. Kuwada, *Convergence of time-inhomogeneous geodesic random walks and its application to coupling methods*, Ann. Probab. **40** (2012), no. 5, 1945–1979. MR-3025706

[26] X.-M. Li, *Effective diffusions with intertwined structures*, arXiv:1204.3250 (2012), 1–33.

[27] X.-M. Li, *Limits of random differential equations on manifolds*, Probab. Theory Related Fields **166** (2016), no. 3-4, 659–712. MR-3568037

[28] X.-M. Li, *Random perturbation to the geodesic equation*, Ann. Probab. **44** (2016), no. 1, 544–566. MR-3456345

[29] M. Liao, *Lévy processes in Lie groups*, Cambridge Tracts in Mathematics, vol. 162, Cambridge University Press, Cambridge, 2004. MR-2060091

[30] O. Lopusanschi and D. Simon, *Area anomaly in the rough path brownian scaling limit of hidden Markov walks*, 2017.

[31] O. Lopusanschi and D. Simon, *Lévy area with a drift as a renormalization limit of Markov chains on periodic graphs*, Stochastic Process. Appl. **128** (2018), no. 7, 2404–2426. MR-3804798

[32] K. V. Mardia and P. E. Jupp, *Directional statistics*, Wiley Series in Probability and Statistics, John Wiley & Sons, Ltd., Chichester, 2000. MR-1828667

[33] M. A. Pinsky, *Isotropic transport process on a Riemannian manifold*, Trans. Amer. Math. Soc. **218** (1976), 353–360. MR-0402957

[34] D. W. Stroock, *Probability theory, an analytic view*, Cambridge University Press, Cambridge, 1993. MR-1267569