International Journal of Mathematics and Mathematical Sciences

WILEY

RESEARCH ARTICLE OPEN ACCESS

On Convergence of a Family of Random Walks in the Infinite Dimensional Stiefel Manifold

Andrea C. G. Mennucci

Scuola Normale Superiore, Piazza dei Cavalieri, 7, Pisa 56126, Italy

Correspondence: Andrea C. G. Mennucci (andrea.mennucci@sns.it)

Received: 3 September 2025 | **Revised:** 16 March 2026 | **Accepted:** 18 March 2026

Academic Editor: Shikha Binwal

Keywords: Brownian motion | Donsker's theorem | Hilbert space | infinite-dimensional manifold | random walk | Riemannian manifold | Stiefel manifold | stochastic completeness | tight family | Wiener process

ABSTRACT

In this paper, we study random walks taking values in an infinite-dimensional space—either a Hilbert space or an infinite-dimensional manifold embedded in it, such as the Stiefel manifold. These random walks arise in problems in shape theory, particularly when stochastic optimization is applied. In these contexts, the random walks are defined at discrete times $t \in \tau = \{t_0 = 0 < t_1 < t_2 < \cdots\}$. By suitably interpolating the paths between times t_i and t_{i+1}, we can view them as time-continuous random walks (for $t \geq 0$). A natural question arises as the fineness of the partition τ tends to zero, does such a family of random walks converge to a stochastic process? This paper provides some preliminary results in this direction, showing that—under appropriate conditions—weak convergence holds, in the sense of Prokhorov's theorem.

MSC2010 Classification: 60J99, 60Fxx, 60B10, 58J65, 58D20, 46T05, 49Q10, 58C35, 60J70, 60B11, 60G50, 60J65, 53A04, 53C22, 58B20, 58D10, 58D15, 58E10

1 | Introduction

1.1 | Shape Theory: Curves

"Shapes" appear in two broad categories of applications.

- Shape optimization, where we want to find the best shape according to a criterion;
- Shape analysis, where we study families of shapes for purposes of statistics, (automatic) cataloging, probabilistic modeling, etc.

Shape theory is central in computer vision because shapes partially characterize objects in images. We focus on the specific case where shapes are represented by smoothly immersed planar curves; this is a widely studied subject (see [1, 2] and references therein). In this case, it should be noted that the shape space classically used in shape optimization is more precisely identified as the space of embedded curves, up to a choice of parameterization, whereas in shape analysis, the space is usually identified as the space of embedded curves, up to rotation, translation, scaling, and reparameterization. We will not address this issue in this paper.

There are various reasons why it is useful to model the space of curves as a Riemannian manifold.

- In the past, methods for shape optimization were devised that would find the solution by using appropriate gradient flows. Calling the minimizing flows gradient flows, however, implies a certain Riemannian metric on the space of curves.

In memory of Prof. Giuseppe Da Prato.

- Modeling the space of curves as a Riemannian manifold also has obvious benefits in shape analysis: Indeed, the distance between curves can be used for clustering, the geodesic can be used to define the average of two shapes, and so on.

We concentrate on two models of "Riemannian manifolds of curves," where we disregard translation and scaling.

- A model for open immersed curves $c: [0, 1] \longrightarrow \mathbb{R}^2$; using a transformation known as square-root velocity representation, the Riemannian manifold is isometric to a subset of the unit sphere in $V = L^2([0, 1])$ (see [3]).
- A model for closed immersed curves $c: S^1 \longrightarrow \mathbb{R}^2$; using an appropriate transformation, the Riemannian manifold is isometric to a subset of the Stiefel manifold $\boldsymbol{St}(2, V)$ (as defined in Section 2.2) (see [4–6]).

Since the sphere in the first model is the special case $\mathbf{St}(1, V)$, we are in both cases interested in infinite dimensional Stiefel manifolds. In both cases, the space of smooth immersions is completed to a larger space of absolutely continuous curves so that the "shape space" is now $\mathbf{St}(p, V)$.

Since stochastic methods play an important role in applications, we are then led to investigate them in $\mathbf{St}(p, V)$. In particular, in [7], a stochastic minimization method was developed to seek numerically global minima for a task of image segmentation; curves would stochastically evolve by a scheme resembling the "random walk on manifold" presented later in Section 6; a pruning method (inspired by simulated annealing) would drive the random walk toward a global minimum (see Figure 1). In [6], a stochastic method was developed in $\mathbf{St}(2, V)$, similar to the classical Kalman filtering, to track a moving object.

A question remained open: Could the numerical methods in [7] and in [6] be explained as a space-time discretization of a yet-to-be-understood infinite-dimensional stochastic method in $\mathbf{St}(2, V)$? Space discretization would not pose a problem since it can be argued that $\mathbf{St}(p, V)$ can be approximated by $\mathbf{St}(p, \mathbb{R}^N)$ for N large (using, *e.g.*, Fourier series). This leaves the following question: Does a discrete time random walk in $\mathbf{St}(p, V)$ somehow approximate a time continuous stochastic process in $\mathbf{St}(p, V)$? More generally, How can we define probabilities and stochastic methods in $\mathbf{St}(p, V)$? Some positive and negative results were found in [8]. In this paper, we will eventually provide a positive result in Theorem 6: The discrete time random walk on $\mathbf{St}(p, V)$ can indeed converge to a time continuous process, when the time partition gets finer and finer. In the spirit of Donsker's theorem, this is a first step to an operative definition of "Brownian motion" on $\mathbf{St}(p, V)$. Accordingly, in the present paper, we focus on tightness and on the existence of limit points. All of the above will eventually provide a sound foundation for methods such as the ones in [7] and in [6].

1.2 | Main Results

The main results of this paper may be summarized as follows. First, in the abstract Hilbert setting, we prove a tightness criterion for interpolated random walks in $C(I; H)$, based on a compact operator K, a first-order approximation of the increments, and suitable growth estimates; this is Theorem 3, and it yields narrow limit points in $C(\mathbb{R}^+; H)$. Second, we show that the same criterion applies to geodesic random walks on manifolds embedded in H, provided the embedding and the exponential map satisfy natural geometric bounds; this is Theorem 5. Third, we verify those hypotheses for infinite-dimensional Stiefel manifolds $\mathbf{St}(p, V)$, obtaining tightness for the corresponding geodesic random walks and an invariance property under the right action of $O(p)$ (see Theorem 6 and the subsequent Corollary). What we do not prove here is uniqueness of the limit, an intrinsic characterization of the limit points, or an identification of the limit with a specific stochastic differential equation; these questions are left for future work.

1.3 | Layout of the Paper

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Let H be a separable Hilbert space and M a manifold embedded in H, possibly infinite dimensional. Let $S = H$ or $S = M$.

Let $\mathfrak{T}$ be a family whose elements τ are discrete countable sets of positive rational numbers (a precise definition will be in Definition 8).

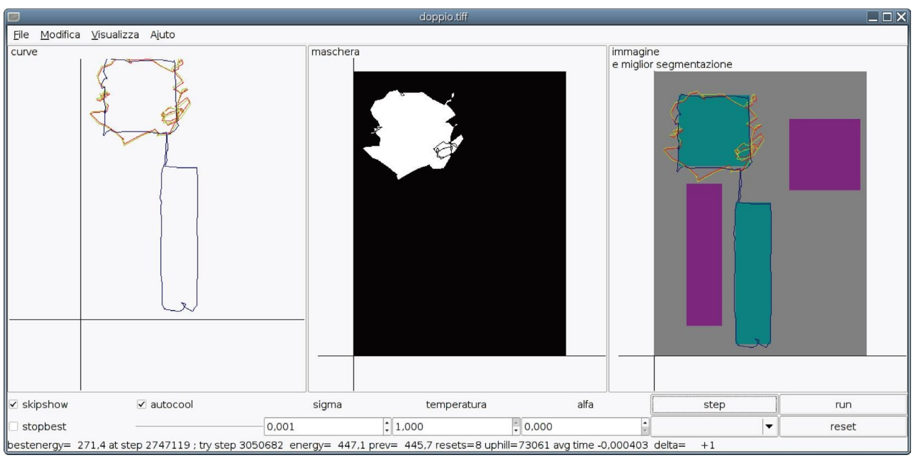

FIGURE 1 | Stochastic minimization of curve-based segmentation energy, from [7]. In the left pane, three curves: the blue curve is the best result so far; in green and red, stochastic steps. In the center pane, the current examined region for segmentation. In the right pane, the image to be segmented, with curves superimposed.

We now describe the organization of the paper and fix the notation for the random walks and the associated interpolated processes.

In this paper, we study processes $\mathfrak{X}^\tau$ which are random functions with values in the Frechét space $C(\mathbb{R}^+; S)$, where $\mathbb{R}^+ = [0, \infty)$. We denote by $\mathfrak{X}^\tau_t \colon \Omega \longrightarrow S$ the random variable obtained from the process $\mathfrak{X}^\tau$ at time t so that $\mathfrak{X}^\tau = (\mathfrak{X}^\tau_t)_{t\in\mathbb{R}^+}$. Each process $\mathfrak{X}^\tau$ is obtained by interpolating a random walk $X^\tau = (X^\tau_n)_{n\in\mathbb{N}}$, observed at discrete times $t \in \tau = \{t_0 = 0 < t_1 < t_2 \dots\}$. More precisely, for $t \in [t_n, t_{n+1}]$, the value $\mathfrak{X}^\tau_t$ is obtained by interpolating between X^τ_n and X^τ_{n+1}.

Section 5 develops an abstract tightness criterion for processes $\mathfrak{X}^\tau$ taking values in a separable Hilbert space $S = H$. Section 6 applies it to geodesic random walks on manifolds $S = M$ embedded in H. Section 7 treats infinite-dimensional Stiefel manifolds and derives the corresponding tightness result. Appendix A contains some useful lemmas, and Appendix B contains deferred proofs.

We will use some definition and results from probability theory; since there may be ambiguity in some definitions, and some results are not easily found in the needed generality, we have written a compendium in [9].

1.4 | Random Walk and Interpolation

Here, we briefly define the discrete-time random walks $X^\tau = (X^\tau_n)_{n\in\mathbb{N}}$ and the associated interpolated processes $\mathfrak{X}^\tau = (\mathfrak{X}^\tau_t)_{t\in\mathbb{R}^+}$ that we will study (more details will be in Section 3.4).

Let $\mathbb{R}^+ = [0, \infty)$; let H, U be separable Hilbert spaces. Let

$$\tau = \{t_0 = 0 < t_1 < t_2, \dots, \} \subset \mathbb{Q}. \tag{1}$$

Note again that these times are assumed to be rational numbers, for convenience (see Lemma 1). We will need a source of random noise: for $t \in \mathbb{R}^+$, a family of i.i.d. r.v. Y_t taking values in U, each with law γ. We will need a Borel map.

$$D = D(x, v, t, s) \colon H \times U \times \left(\mathbb{R}^+\right)^2 \longrightarrow H, \tag{2}$$

continuous in s and such that $D(x, v, t, 0) = 0$. We fix $X^\tau_0 = \mathfrak{X}_0$ a random variable, and we define recursively

$$X^\tau_{n+1} = X^\tau_n + D\big(X^\tau_n, Y_{t_n}, t_n, (t_{n+1} - t_n)\big). \tag{3}$$

Then, we interpolate using

$$\mathfrak{X}^\tau_t = X^\tau_n + D\big(X^\tau_n, Y_{t_n}, t_n, (t - t_n)\big), \tag{4}$$

for $t_n \le t \le t_{n+1}$; so, each trajectory $t \longmapsto \mathfrak{X}^\tau_t(\omega, t)$ is continuous; hence, each $\mathfrak{X}^\tau$ is a r.v. taking value in $C(\mathbb{R}^+; H)$, the Frechét space of continuous functions $x \colon \mathbb{R}^+ \longrightarrow H$.

Since U is used only in the second argument of D, and H and U are isomorphic, we can decide that $H = U$ with no loss of generality.

2 | Motivation

2.1 | Finite Dimensional Manifolds

The theory of stochastic differential equations in finite dimensional Riemannian manifolds M is well developed; see, e.g., [10]. In particular, there are multiple equivalent definitions of Brownian motion, each based on different principles.

- Stochastic differential equations in local charts,
- Development of Euclidean Brownian motion (Example 2.6.8 in [10]),
- The heat equation and its transition probabilities,
- Limit of geodesic random walks[1],

but all leading to the same ultimate definition; see Proposition 3.2.1 in [10].

We define this concept as in Section 4.2 in [10].

Definition 1 (Stochastic completeness). We consider a noncompact connected manifold M, and let ∞ be the point added by the Alexandroff compactification (the one-point topological compactification). For any continuous path $x \colon \mathbb{R}^+ \longrightarrow M \cup \{\infty\}$, let

$$e = e(x) = \sup\{t \ge 0 \colon \forall s, 0 \le s < t, x(s) \in M\}, \tag{5}$$

be the first time t such that $x(t) = \infty$; we agree that $x(s) = \infty$ for $s \ge e$. Suppose that $\mathfrak{X} = (\mathfrak{X}_t)_{t\in\mathbb{R}^+}$ is Brownian motion, whose paths are continuous in $M \cup \{\infty\}$. A manifold is called stochastically complete if $e(\mathfrak{X}_\cdot)$ is infinite almost surely

$$\mathbb{P}\{e(\mathfrak{X}_\cdot) = \infty\} = 1. \tag{6}$$

A thorough discussion of this problematic may be found in [11]. There is an important problem: Even if the manifold is complete, it may fail to be stochastically complete ([11] attributes the first such example to [12]).

There are many properties of M that ensure that the manifold is stochastically complete, such as volume growth of geodesic balls, isoperimetric inequalities, conservation of mass in the heat equation, and curvature bounds; see [11] (indeed, our Theorem 5 requires a kind of curvature bound).

2.1.1 | Infinite Dimensional Manifolds

When the Riemannian manifold M is infinite dimensional, though, we immediately identify some obstacles.

- When the manifold M has dimension N, we have an important property: Each tangent space T_xM is isomorphic to $\mathbb{R}^N$; hence, there is a canonical choice of Gaussian measure $N(0, \mathbb{I})$ on each one. This is, in a sense, the "white noise" that is driving the Brownian motion.

  When the manifold M is modeled on a infinite dimensional Hilbert space H, then there is no Gaussian measure in H that is rotationally invariant (actually, rotations of a Gaussian measure $N(0, Q)$ tend to be mutually singular, as explained in [13]). So, we will need to decide what "white noise" we will use.

- While the heat equation can be defined in H, an approach using this tool would have to deal with some technical difficulties; for example, the heat equation is not Feller; that is, it does not regularize the initial data.

  Moreover, usually the transition probabilities of the heat kernel are used; these transition probabilities are expressed as densities with respect to the volume form, but an infinite dimensional Riemannian manifold does not have a volume form that may be used as a reference measure.

- The Hopf–Rinow theorem is false; closed bounded sets are not necessarily compact.
- The one point compactification is not useful, since any nonempty open set contains a sequence such that $x_n \longrightarrow \infty$.
- The radial process is not useful since there may be examples of complete Riemannian manifolds where the trajectories of the Brownian motion are bounded, but each of them would have $\mathfrak{X}_t \longrightarrow \infty$ in finite time (a key point to build such an example may be [14]).

Prokhorov's theorem, on the other end, is valid in any separable metric space (regardless of "dimension"); so, an appropriate concept of tightness will be the key element for Theorem 5.

For all the reasons above, it may be quite difficult to refer to preexisting studies and results of Brownian motions on finite-dimensional manifolds, when addressing the problem here presented.

2.2 | Stiefel Manifolds

The results in this paper will be valid when M is a Stiefel manifold.

Let n and p be natural numbers with $1 \le p \le n$; classically, the Stiefel manifold $\boldsymbol{St}(p, \mathbb{R}^n)$ is defined as the set of all frames composed of p orthonormal vectors in $\mathbb{R}^n$; those frames are represented as $n \times p$ matrices. Geodesics in Stiefel manifolds $\boldsymbol{St}(p, \mathbb{R}^n)$ are known to have closed form solutions as demonstrated by Edelman et al. [15] in Section 2.2.2. This property extends to infinite dimensional manifolds, as will be explained in Section 7.3.1.

Let again $p \ge 1$ be a natural number, and let V be a Hilbert space with $\dim(V) \ge 2p$ (possibly infinite dimensional). Let $H = V^p$, then we write

$$x \in H, x = (x_1, \ \dots x_p), \tag{7}$$

and

$$\|x\|_H = \sqrt{\sum_{i=1}^{p} \|x_i\|_V^2}, \tag{8}$$

as usual (and similarly for scalar products). By analogy to the finite dimensional space, we will call columns the p vectors x_i that compose x. We define

$$\pi_i : H \longrightarrow V, \pi_i(x) = x_i. \tag{9}$$

Definition 2. We define $\mathbf{St}(p, V)$ as the manifold of $x \in H$ such that

$$\langle x_i, x_j \rangle_V = \delta_{i,j}, \tag{10}$$

and $\mathbf{St}(p, V)$ is an embedded manifold in H of codimension $p(p+1)/2$.

The sphere is the special case $\mathbf{St}(1, V)$. The geometry of Stiefel manifolds is pretty well understood [15, 16] (see Section 7.3.1 for details).

3 | Definitions

Let H be a separable Hilbert space. We agree that variables n, m, h, k, i, j, l are natural numbers. We will use the theory of *"nets"*; in this paper, a net is a function whose domain is a partially ordered directed set with no maxima (abbreviated to "dposet" in the following; see [9] for details and properties).

3.1 | Measures

Definition 3. Given measurable spaces (X_1, Σ_1) and (X_2, Σ_2), a measurable mapping $f: X_1 \longrightarrow X_2$, and a measure $\mu: \Sigma_1 \longrightarrow [0, +\infty]$, the push forward of μ is defined to be the measure $f_\sharp\mu: \Sigma_2 \longrightarrow [0, +\infty]$ given by

$$f_\sharp\mu(B) = \mu\big(f^{-1}(B)\big), \quad \text{for all}\, B \in \Sigma_2. \tag{11}$$

The push forward measure is denoted also as $f_*\mu$, $\mu \circ f^{-1}$, or $f \,\#\, \mu$.

Definition 4 (Law *a.k.a.* distribution). If $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space and $Y: \Omega \longrightarrow X_2$ is a r.v. and γ is a probability measure on (X_2, Σ_2), we will write

$$Y \sim \gamma \,\text{when}\, \gamma = Y_\sharp\mathbb{P}, \tag{12}$$

we will say that γ is the law or the distribution of Y; similarly, for $Y, Z: \Omega \longrightarrow X_2$,

$$Y \sim Z \,\text{when}\, Y_\sharp\mathbb{P} = Z_\sharp\mathbb{P}. \tag{13}$$

We will use the narrow convergence.[3]

Definition 5 (Narrow convergence). Let S be a Hausdorff topological space. Given a net of Radon measures μ_α, μ on S, for $\alpha \in A$ a dposet, we will say that $\mu_\alpha \longrightarrow \mu$ narrowly if

$$\forall f \in C_b(S), \lim_\alpha \int_S f(x) \mathrm{d}\mu_\alpha(x) = \int_S f(x) \mathrm{d}\mu(x). \tag{14}$$

We agree on this (nonstandard) definition.

Definition 6. Let $\mathfrak{T}$ be a dposet and S be a Hausdorff topological space. Let μ_τ be a net of Radon measures on S: It is tight that if $\forall \varepsilon > 0$, there is a compact set $C \subseteq S$ such that

$$\limsup_{\tau \in \mathfrak{T}} \mu_\tau(S \backslash C) \le \varepsilon. \tag{15}$$

Let $\mathfrak{X}^\tau$ be a net of r.v. taking values in S, for $\tau \in \mathfrak{T}$: It is **tight** if the net of laws $\mu_\tau = \mathfrak{X}^\tau_\sharp\mathbb{P}$ is tight.

3.2 | Probability Setting

Hypotheses 1. We fix a constant $c_t > 0$ that will be used to bound temporal finess and a constant $c_3 > 0$ that will be used to control exponential decay.[4]

We will use a Borel measure γ on H satisfying

-

$$\int_H \|x\|^4 e^{4c_3 c_t \|x\|} \mathrm{d}\gamma(x) < \infty. \tag{16}$$

- γ is centered

$$\int_H x \, d\gamma(x) = 0; \tag{17}$$

- There is [5] a linear compact symmetric injective operator $K: H \longrightarrow H$ such that $\gamma(K(H)) = 1$ and

$$\int_H \left\|K^{-1}x\right\|^4 d\gamma(x) < \infty. \tag{18}$$

All the above hold true when $\gamma = N(0, Q)$ a Gaussian measure (as defined in next section); see Proposition 1 and Remark 4.

Let γ be such a probability on H. We need a probability space so that we have i.i.d. r.v. $Y_t: \Omega \longrightarrow H$ each with distribution $Y_{t\sharp}\mathbb{P} = \gamma$, for $t \in \mathbb{Q}$; to this end, we may set

$$\Omega = \times_{t\in\mathbb{Q}} H, \mathcal{F}, \mathbb{P} = \otimes_{t\in\mathbb{Q}}\gamma; \tag{19}$$

where $\mathcal{F}$ is the σ-algebra generated by null sets of $\mathbb{P}$ and by $\mathcal{B}(\Omega)$, the Borel σ-algebra.

Proposition 1. *For any $c \in [0, 4c_3]$, $\alpha \in [0, 4]$, there is a constant $\widetilde{c} = \widetilde{c}(\alpha, c, c_t) > 0$ such that for all $\delta \in [0, c_t]$,*

$$\mathbb{E}[g(\delta\|Y_t\|)] \le \widetilde{c}\delta^\alpha, \tag{20}$$

where $g(s) = s^\alpha e^{cs}$ (the proof is in Section B.)

3.3 | Gaussian Measures

Let H be a separable Hilbert space.

Definition 7. Suppose that $a \in H$ and $Q: H \longrightarrow H$ is a linear symmetric trace-class operator such that the quadratic form $\langle x, Qx\rangle_H$ is nonnegative. We recall that $\gamma = N(a, Q)$ is a Gaussian measure in the Hilbert space H when the characteristic function (or Fourier transform) is

$$\forall f \in H, \int_H e^{i\langle f,x\rangle} d\gamma(x) = \exp\Big(i\langle a, f\rangle - \frac{1}{2}\langle x, Qx\rangle_H\Big). \tag{21}$$

(Theorem 2.3.1 in [13]; Section 1.5 in [17]) The mean and variance are characterized by

$$\langle f, a\rangle = \int_H \langle f, x\rangle d\gamma(x), \tag{22}$$

$$\langle f, Qg\rangle_H = \int_H \langle f, x - a\rangle\langle g, x - a\rangle d\gamma(x), \tag{23}$$

for all $f, g \in H$. In particular, γ is called centered when $a = 0$ and nondegenerate when the variance has empty kernel.

The Proposition 1 is true for Gaussian measures.

Proposition 2. *Let $c \ge 0, \alpha > 0$ and let $g(s) = s^\alpha e^{cs}$; suppose $Y \sim N(0, Q)$, then there is a $\widetilde{c} = \widetilde{c}(\alpha, c, Q, c_t) > 0$ such that for $0 \le \delta \le c_t$,*

$$\mathbb{E}[g(\delta\|Y\|)] \le \widetilde{c}\delta^\alpha. \tag{24}$$

The proof is in Section B.

3.4 | Random Walk

As aforementioned, in this paper, a net is a function whose domain is a partially ordered directed set with no maxima (abbreviated to "dposet"). We will use these dposets.

Definition 8. We fix a constant $c_t > 0$, the same constant as in Hypothesis 1.

1. Let

$$\tau = \{t_0 = 0 < t_1 < t_2, \dots,\} \subset \mathbb{Q} \tag{25}$$

be such that

$$\lim_{n\longrightarrow\infty} t_n = \infty, \sup_n(t_{n+1} - t_n) \le c_t. \tag{26}$$

Let $\mathfrak{T}$ be the dposet of all such τ, ordered by inclusion.

2. Let $T > 0$ we define $\mathfrak{T}_T$ be the dposet of all τ of the form

$$\tau = \{t_0 = 0 < t_1 < t_2 <, \dots, < t_n = T\}, \tag{27}$$

with $t_0, t_1, \dots t_{n-1} \in \mathbb{Q}$ and again $\max_{1\le j\le n}(t_j - t_{j-1}) \le c_t$ (note that we do not require that $T \in \mathbb{Q}$.)

We will use $\mathfrak{T}$ for processes with $t \in \mathbb{R}^+ = [0, \infty)$ while we will use $\mathfrak{T}_T$ for processes with $t \in [0, T]$. We will actually define all processes as in the first case, for simplicity; but then, up to restricting $t \in [0, T]$, we will study tightness using $\tau \in \mathfrak{T}_T$.

Let H be a separable Hilbert space.

Definition 9. We will need a Borel map

$$D: H^2 \times \left(\mathbb{R}^+\right)^2 \longrightarrow H, \tag{28}$$

continuous in the last argument and such that $D(x, v, t, 0) = 0$.

For each τ, we define an interpolated process $\mathfrak{X}^\tau = (\mathfrak{X}^\tau_t)_{t\ge0}$ taking values in H.

We fix $\mathfrak{X}_0$ a random variable taking values in H, independent of all Y_t.

To define $\mathfrak{X}^\tau$, we define auxiliary random walks X^τ_n for $n \in \mathbb{N}$, where we define $X^\tau_0 = \mathfrak{X}_0$, and we define recursively

$$X^\tau_{n+1} = X^\tau_n + D\big(X^\tau_n, Y_{t_n}, t_n, (t_{n+1} - t_n)\big), \tag{29}$$

then we interpolate using

$$\mathfrak{X}^\tau_t = X^\tau_n + D\big(X^\tau_n, Y_{t_n}, t_n, (t - t_n)\big), \tag{30}$$

for $t_n \le t \le t_{n+1}$; so, each trajectory $t \longmapsto \mathfrak{X}^\tau_t(\omega, t)$ is continuous; hence, each $\mathfrak{X}^\tau$ is a r.v. taking value in $C(\mathbb{R}^+; H)$, the Fréchet space of continuous functions $x: \mathbb{R}^+ \longrightarrow H$ with $\mathbb{R}^+ = [0, \infty)$.

Note that, for all τ, we have $\mathfrak{X}_0 = \mathfrak{X}^\tau_0$.

Proposition 3 (Filtration). *Let $\mathcal{F}_t$ be the σ-algebra generated by $\mathfrak{X}_0$ and by Y_s for $s < t$ (augmented with the null sets of $\mathbb{P}$). The process X^τ_n for $t_n \le t$ is $\mathcal{F}_t$-measurable. The process $\mathfrak{X}^\tau_s$ for $s \le t$ is $\mathcal{F}_t$-measurable.*

Remark 1. The choice of interpolation (4) has a beneficial effect: suppose T is positive but $T \notin \tau$; define $\hat{\tau} = (\tau \cup \{T\}) \cap [0, T]$ so that $\hat{\tau} \in \mathfrak{T}_T$, then

$$\mathfrak{X}_t^{\hat{\tau}} = \mathfrak{X}_t^{\tau} \quad \forall t \le T. \tag{31}$$

This means that, up to adding T to τ, we can consider any process defined above as a process $\mathfrak{X}_t^\tau$ for $t \in [0, T]$ and $\tau \in \mathfrak{T}_T$ (note that we do not require that $T \in \mathbb{Q}$).

We recall that $C(I; H)$ is a complete separable metric space; so, when the topology associated to probability measures on $C(I; H)$ is the narrow topology and the family is tight (as defined in Definition 6), by Prokhorov's theorem[6], the set of limit points is not empty. When $I = [0, T]$, this can also be explained using sequences.

Lemma 1. *Let $T > 0, I = [0, T]$; let $q_0 = 0, q_1 = T$ and*

$$\{q_2, q_3, \dots\} = (0, T) \cap \mathbb{Q}, \tag{32}$$

be an enumeration, and let

$$\theta_n = \{q_0, q_1, \dots, q_n\}; \tag{33}$$

then, for $n \ge 2$, the sequence $(\theta_n)_n$ is cofinal in $\mathfrak{T}_T$. So, the limit points

$$\bigcap_{\hat{\tau} \in \mathfrak{T}_T} \overline{\{\mathfrak{X}^\tau : \hat{\tau} \subseteq \tau\}}, \tag{34}$$

along with $\mathfrak{T}_T$ coincide with the limit points

$$\bigcap_{n \in \mathbb{N}} \overline{\{\mathfrak{X}^{\theta_k} : k \ge n\}}, \tag{35}$$

along the sequence $(\theta_n)_n$. Similarly, limits, limsup, liminf, tightness, etc., can be defined using that sequence $(\theta_n)_n$.

A similar result does not hold for $\mathfrak{T}$.

Proposition 4. *There does not exist a cofinal sequence $f: \mathbb{N} \longrightarrow \mathfrak{T}$ (the proof is in Appendix B).*

3.5 | Manifold

Suppose M is a manifold smoothly embedded in H; we consider it as a Riemannian manifold since it inherits the scalar product from H. For $x \in M, v \in T_x M$, we denote by $\exp_x(v)$ the exponential map. We require that M be a closed subset, so it is geodetically complete. We consider $T_x M$ as a linear subspace of H, not as its affine translation containing x. For $x \in M$, we define the orthogonal projection $\pi_{T_x M}: H \longrightarrow T_x M$; note that $\pi_{T_x M}$ is symmetric, that is, $\pi_{T_x M} = \pi^*_{T_x M}$; we will call $P_x = \pi_{T_x M}$ for simplicity.

3.5.1 | Geodesic Random Walks on Manifolds

In this case, for each τ, we define an interpolated process $\mathfrak{X}^\tau = (\mathfrak{X}_t^\tau)_{t \ge 0}$ taking values in M. We fix $\mathfrak{X}_0$ a random variable taking values in M, independent of all Y_t. We define recursively $X_0^\tau = \mathfrak{X}_0$ and

$$X_{n+1}^\tau = \exp_{X_n^\tau}\left(\sqrt{(t_{n+1} - t_n)}\, P_{X_n^\tau} Y_{t_n}\right). \tag{36}$$

Then again, we define $\mathfrak{X}^\tau$ by interpolating along geodesics

$$\mathfrak{X}_t^\tau = \exp_{X_n^\tau}\left(\sqrt{(t - t_n)}\, P_{X_n^\tau} Y_{t_n}\right), \tag{37}$$

for $t_n \le t \le t_{n+1}$, and again, $t \longmapsto \mathfrak{X}_t^\tau$ is continuous.

This kind of random walk will be called geodesic random walk.

If, for $x \in M, y \in H$, we let

$$D(x, y, t, s) \overset{\text{def}}{=} \exp_x\left(\sqrt{s}\, P_x y\right) - x, \tag{38}$$

then we obtain the same definition as in the previous Section 3.4; so all comments and results apply to this case as well. Note that, in the words of Section 2.1.1, $P_x Y_t$ is the source of "white noise" that we are using to drive the random walk. We will come back to manifolds in Section 6 and to Stiefel manifolds in Section 7

4 | Tightness by Ascoli–Arzelà Theorem

In the following, for $\psi: \mathbb{R} \longrightarrow \mathbb{R}$, "monotonic" means monotonically weakly increasing, that is, $s \le t \Longrightarrow \psi(s) \le \psi(t)$.

Definition 10. Let $I \subseteq \mathbb{R}$ be an interval and E be a normed vector space with norm $\|\cdot\|_E$; for $x: I \longrightarrow E$ uniformly continuous and $\eta > 0$, we define the modulus of continuity

$$\omega_{I,E}(x, \eta) \overset{\text{def}}{=} \sup\{\|x(t) - x(s)\|_E : t, s \in I, |t - s| \le \eta\}; \tag{39}$$

note that $\omega_{I,E}(x, \cdot)$ is continuous, subadditive, monotonic, and $\omega_{I,E}(x, 0) = 0$.

We recall that a set is called precompact if its closure is compact. We will use this version of Ascoli–Arzelà theorem (recall that if I is compact then $C(I; S) = C_b(I; S)$)

Theorem 1. *Suppose H is a Banach space. Let $I \subseteq \mathbb{R}$ be a compact interval. Let $F \subseteq C(I; H)$ be a family of continuous functions $x: I \longrightarrow H$. We consider the following two clauses:*

- *There is $J \subseteq I$ countable dense subset such that for each $t \in J$, there exists a precompact set $C_t \subset H$ such that $\forall x \in F, x(t) \in C_t$;*
-

$$\lim_{\eta \longrightarrow 0} \sup_{x \in F} \omega_{I,H}(x, \eta) = 0. \tag{40}$$

The above two clauses hold if and only if F is precompact in $C(I; H)$.

This is a Banach-valued version of the classical Ascoli–Arzelà criterion. For the reader's convenience, the idea is as follows: one first uses clause (i) and a diagonal argument on the countable dense set J to extract a subsequence converging pointwise on J; clause (ii) gives a uniform modulus of continuity, which extends the pointwise limit on J to a continuous limit on I and upgrades convergence on J to uniform convergence on all of I. The converse implication is immediate from precompactness in $C(I; H)$.

For probability theory, we transform the above as follows.

Theorem 2. *Suppose H is a Banach space. Let $I = [0, T]$. Suppose that $\mathfrak{X}^\alpha: \Omega \longrightarrow C(I; H)$ is a net of processes (with $\alpha \in A$) satisfying the following two clauses.*

- $\forall \varepsilon > 0$, *there exist a countable set* $J = \{a_0, a_1, \dots a_j \dots\}$ *dense in* I, *and compact sets* $C_j \subset H$, *such that*

$$\forall \alpha \in A\, \forall j \in \mathbb{N}, \mathbb{P}\{\mathfrak{X}^\alpha(a_j) \notin C_j\} \le \varepsilon 2^{-j}. \tag{41}$$

- $\forall \varepsilon_0 > 0$, $\forall \varepsilon_1 > 0$, *and* $\exists \eta > 0$, $\exists \alpha_0 \in A$ *such that*

$$\forall \alpha \in A, \alpha \ge \alpha_0 \Longrightarrow \mathbb{P}\{\omega_{I,H}(\mathfrak{X}^\alpha, \eta) \ge \varepsilon_0\} \le \varepsilon_1. \tag{42}$$

Then, the net $\mathfrak{X}^\alpha$ *is tight*[7] *in* $C(I; H)$.

This follows immediately from Theorem 1; for a more expanded version in the same setting, see Theorem 6.28 in [9].

Remark 2. The second hypothesis (42) may be reformulated as

$$\forall \varepsilon_0 > 0, \lim_{\eta \longrightarrow 0} \limsup_{\alpha} \mathbb{P}\{\omega_H(I, \mathfrak{X}^\alpha, \eta) \ge \varepsilon_0\} = 0, \tag{43}$$

since ω is monotonic in η.

5 | Tightness of Random Walks

Let again H be a separable Hilbert space. The purpose of this section is to state and prove Theorem 3 on tightness of random walks in H.

5.1 | Tightness Operator

Let γ be as defined in Hypotheses 1. We required in Hypotheses 1 that there is $K: H \longrightarrow H$ a linear symmetric compact injective operator such that $\gamma(K(H)) = 1$, and equivalently,

$$\mathbb{P}(Y_t \in K(H)) = 1\, \forall t. \tag{44}$$

We define now

$$D^K(0, r) = \{Kx: x \in H, \|x\| \le r\} = K(B_H(0, r)) = rK(B_H(0, 1)), \tag{45}$$

then

- They are precompact (this means that the closures $\overline{D^K(0, r)}$ are compact), and
- $$\bigcup_n D^K(0, n) = K(H), \tag{46}$$

  so

  $$\gamma\Big(\bigcup_n D^K(0, n)\Big) = 1; \tag{47}$$

- this means that $\forall \varepsilon > 0\ \exists\ n$ such that

$$\gamma(D^K(0, n)) \ge 1 - \varepsilon. \tag{48}$$

Remark 3. Since H is a separable Hilbert space, then every probability measure on it is Radon hence tight; the above gives us an accessible family of sets for tightness of γ.

Up to rescaling K, we will assume that

$$\forall v \in H, \|Kv\|_H \le \|v\|_H. \tag{49}$$

Remark 4. In general, for any Hilbert space H and Gaussian measure $\gamma = N(0, Q)$, such operator K always exists.

Remark 5. The role of the compact operator K is to provide a compact scale on the Hilbert space H. In infinite dimension, bounded sets in the ambient norm $\|\cdot\|_H$ are typically not compact, so moment bounds in H alone do not yield tightness. By contrast, if K is a compact operator, the sets $D^K(0, r)$ are precompact in H; equivalently, boundedness in the K-norm means confinement in a precompact subset of H. The first part of the proof of Theorem 3 uses K to obtain compact containment at fixed times, while the second part uses the ambient H-norm to control the modulus of continuity. Gaussian measures naturally provide such a compact scale, but the theorem is formulated abstractly and applies whenever a suitable compact injective operator K is available for the law of the source of random noise Y_t.

5.2 | Tightness Theorem

Definition 11. Given a linear continuous injective operator $K: H \longrightarrow H$, we define

$$\|v\|_K \stackrel{def}{=} \begin{cases} \|K^{-1}v\|_H, & \text{if } v \in K(H), \\ \infty, & \text{if } v \notin K(H). \end{cases} \tag{50}$$

Similarly, for scalar products, for $v, w \in K(H)$, we define

$$\langle v, w\rangle_K = \langle K^{-1}v, K^{-1}w\rangle_K. \tag{51}$$

Definition 12. Given Banach spaces B_1 and B_2, we define $\mathcal{L}(B_1; B_2)$ to be the space of linear continuous operators $A: B_1 \longrightarrow B_2$. If $B_1 = B_2$, then we write $\mathcal{L}(B_1)$.

Theorem 3. *Let* $I = [0, T]$. *Suppose that the processes* $\mathfrak{X}^\tau: \Omega \longrightarrow C(I; H)$ *above defined in Definition 9 satisfy the following clauses.*

1. *There is a linear compact operator* $K: H \longrightarrow H$ *satisfying the requisites in the previous section and such that*

$$\mathbb{E}\big[\|\mathfrak{X}_0\|_K^4\big] < \infty. \tag{52}$$

2. *For all* $x, v \in K(H), t \in I, s \ge 0$,

$$D(x, v, t, s) \in K(H). \tag{53}$$

3. *There is a bounded Borel functional* $L = L(x, t)$

$$L: H \times \mathbb{R}^+ \longrightarrow \mathcal{L} = \mathcal{L}(H; H), \tag{54}$$

such that $\forall x, v \in K(H)$

$$L(x, t)v \in K(H); \tag{55}$$

4. *There are constants* $c_3, c > 0$ *such that for all* $x, v \in K(H)$,

$$\|L(x,t)v\|_K \le c\big(\|v\|_K + \|x\|_K\|v\|_H\big)e^{c_3\|v\|_H}, \tag{56}$$

(the constant c_3 must be the same as in Hypotheses 1).

5. *Moreover, for all $x, v \in K(H)$, $t \in I$, and $s \in [0,1]$,*

$$\|D(x,v,t,s) - \sqrt{s}\,L(x,t)v\|_K \le cs\big(\|x\|_K\|v\|_H^2 + \|v\|_K\|v\|_H\big)e^{c_3\sqrt{s}\|v\|_H}. \tag{57}$$

6. *There is a constant $c_d > 1$ such that $\forall x, v \in H, \forall t \in \mathbb{R}^+$, $\forall s \in [0,1]$,*

$$\|D(x,v,t,s)\|_H \le c_d\sqrt{s}\|v\|_H, \tag{58}$$

$$\sqrt{s}\|v\| \le 1/c_d \Longrightarrow \|D(x,v,t,s) - \sqrt{s}\,L(x,t)v\|_H \le c_d s\|v\|_H^2. \tag{59}$$

Then, the family $\mathfrak{X}^\tau$, for $\tau \in \mathfrak{T}_T$, is tight in $C(I;H)$.

The assumptions of Theorem 3 play two different roles. Item 1 introduces the compact scale K and controls the initial condition in that scale. Items 2, 3, 4, and 5 ensure that the discrete-time random walk and its first-order approximation are compatible with the K-structure so that one can obtain compact-containment estimates at fixed times. Item 6, instead, gives coarser estimates in the ambient Hilbert norm H; these are used to control increments and the modulus of continuity, which is the second ingredient in the tightness argument. In this sense, the theorem is an abstract criterion: The first group of assumptions provides compactness, while the last one provides pathwise regularity. To prove this theorem, we apply Theorem 2 and verify its two hypotheses separately. The proof is in Section 5.4.

Remark 6. In particular, $D(x,v,t,s)$ is Frechét differentiable in v at $v=0$, and the differential is $\sqrt{s}L(x,t)$, and by equation (58), we have

$$\|L(x,t)\|_{\mathcal{L}} \le c_d, \quad \forall x \in H. \tag{60}$$

Remark 7. The third assumption in Hypotheses 1 and the first two hypotheses above imply that

$$\forall t \ge 0, \mathbb{P}\{\mathfrak{X}_t^\tau \in K(H)\} = 1. \tag{61}$$

5.2.1 | Corollaries

Corollary 1. *Suppose that the hypotheses of Theorem 3 hold for any $T > 0$; then, the family of processes $\mathfrak{X}^\tau$, for $\tau \in \mathfrak{T}$, is tight in $C(\mathbb{R}^+;H)$.*

Proof 1. We recall this fact. Let $T > 0$; we consider the restriction map

$$r_T\colon C(\mathbb{R}^+;H) \longrightarrow C([0,T];H), \tag{62}$$

given by $r_T f = f_{|[0,T]}$; then, the topology on $C(\mathbb{R}^+;H)$ is the initial topology with respect to the maps r_n and the Banach spaces $C([0,n];H)$, for $n \in \mathbb{N}$ (see also Section 5 in [9] for more details).

Let $\varepsilon > 0$, for any $n \in \mathbb{N}$, by Theorem 3, there exists a compact set $E_n \subseteq C([0,n];H)$ such that

$$\mathbb{P}\{\mathfrak{X}^\tau \notin r_n^{-1}(E_n)\} \le \varepsilon 2^{-n}, \tag{63}$$

and let

$$E = \{f \in C(\mathbb{R}^+;H)\colon \forall n \in \mathbb{N}, r_n f \in E_n\}, \tag{64}$$

then (by a diagonal argument), E is precompact in $C(\mathbb{R}^+;H)$ and

$$\begin{aligned}\mathbb{P}\{\mathfrak{X}^\tau \notin E\} &= \mathbb{P}\{\mathfrak{X}^\tau \in E^c\} = \mathbb{P}\Big\{\mathfrak{X}^\tau \in \bigcup_n r_n^{-1}(E_n^c)\Big\}\\ &\le \sum_n \mathbb{P}\{\mathfrak{X}^\tau \in r_n^{-1}(E_n^c)\} = \sum_n \mathbb{P}\{r_n \circ \mathfrak{X}^\tau \notin E_n\} \le 2\varepsilon.\end{aligned} \tag{65}$$

□

More detailed and more general statements can be found in Section 6.3 in [9].

Since $C(\mathbb{R}^+;H)$ is a Fréchet space, by Prokhorov's theorem, we obtain this result.

Corollary 2. *Suppose that the hypotheses of Theorem 3 hold for any $T > 0$; the net of processes $\mathfrak{X}^\tau$, as r.v. in the Frechét space $C(\mathbb{R}^+;H)$, has narrow limit points.*

Corollary 3. *By Lemma 4, we have*

$$\mathbb{E}\Big[\,\|\mathfrak{X}_t^\tau\|_K^4\Big] < \infty, \tag{66}$$

so the process $\mathfrak{X}_t^\tau$ can be restarted at time t using $\mathfrak{X}_t^\tau$ as initial time.

Suppose that the family $\mathfrak{X}^\tau$ is tight in $C(I;H)$ (with $I = [0,T]$ or $I = [0,\infty)$), then by Prokhorov's theorem, the set of limit points is not empty; obviously, being $\mathfrak{X}$ a random variable $C(I;H)$, then (almost all) paths are continuous. Something more can be said.

Corollary 4. *Any limit point $\mathfrak{X}$ will have a version such that almost all trajectories $t \longmapsto \mathfrak{X}_t$ of $\mathfrak{X}$ are Hölder continuous functions with an arbitrary exponent smaller than 1/4.*

Proof 2. We use Theorem 6.19 from [9] with $p_1 < p_2 = 4$ and use Lemma 4 in the following to state that

$$\lim_\tau \mathbb{E}\Big[\,\|\mathfrak{X}_t^\tau - \mathfrak{X}_s^\tau\|_H^{p_1}\Big] = \mathbb{E}\big[\|\mathfrak{X}_t - \mathfrak{X}_s\|_H^{p_1}\big] \le c|t-s|^{p_1/2}. \tag{67}$$

We use the Kolmogorov test[8]: we apply it with $\delta = p_1$, $\varepsilon = 1$ and replacing

$$\rho(Z(t),Z(s))^\delta = \|\mathfrak{X}_t - \mathfrak{X}_s\|_H^{p_1}, \tag{68}$$

so there is a version where paths are Hölder continuous functions with an arbitrary exponent smaller than $1/2 - 1/p_1$. □

Remark 8. At this level of generality, we do not expect that there is a unique limit point.

5.3 | A Basic Example

We recall this standard result.

Theorem 4 (Donsker's theorem). *Let $\xi_1, \xi_2, \xi_3, \ldots$ be a sequence of i.i.d. real random variables with mean 0 and variance 1. Let*

$$S_n = \sum_{i=1}^{n} \xi_i, \tag{69}$$

and $S_0 = 0$. The stochastic process S_n is known as a random walk.

We rescale and extend the process $(S_n)_n$ to continuous time $t \in [0, 1]$. We define

$$W^N(t) = \frac{S_j}{\sqrt{N}}, \quad t = \frac{j}{N}, j = 0, \dots, N, \tag{70}$$

and then linearly interpolate

$$W^N(t) = (1-s)\frac{S_j}{\sqrt{N}} + s\frac{S_{j+1}}{\sqrt{N}} = \frac{S_j + s\xi_{j+1}}{\sqrt{N}}, \tag{71}$$

for $j/N \le t \le (j+1)/N$ and

$$t = (1-s)\frac{j}{N} + s\frac{j+1}{N} = \frac{j}{N} + s\frac{1}{N}, \quad \text{i.e.} s = Nt - j \in [0,1]; \tag{72}$$

so, W^N is a random variable taking values in $C([0, 1])$.

Then, the sequence of random function $(W^N)_N$ converges narrowly to a random function W on $C([0, 1])$, as $N \longrightarrow \infty$; this W is the standard Wiener Process 9.

Theorem 9.1 in [18] uses the above as a method to define the Wiener process; by Kolmogorov's theorem, W has a version where almost all paths are continuous. Other sources call the above result Donsker's theorem.

The random walk $(S_n)_n$ can be recast into a continuous time process in different ways.

- We may define

$$W^N(t) = \frac{S_{\lfloor Nt \rfloor}}{\sqrt{N}}, \quad t \in [0,1], \tag{73}$$

  and consider W^N as random variables taking values in the Skorokhod space $\mathbb{D}[0, 1]$.

- We may define

$$W^N(t) = \frac{S_j + \sqrt{s}\xi_{j+1}}{\sqrt{N}}m, \tag{74}$$

  for t, s defined as in equation (72).

With those (and other) interpolations, the thesis of Theorem 4 holds as well.

Note that the interpolation (74) has a benefit: if the variables $\xi_j \sim N(0, 1)$, then $W^N(t) \sim N(0, t)$.

Remark 9. The above construction of W^N in Theorem 4, when using the interpolation (74) is a special case of the random walk W^N where $\tau^N = \{t_j = j/N : j = 0, \dots, N\}$, $\mathfrak{X}_0 = 0$, $H = \mathbb{R}$, and $D(x, v, t, s) = \sqrt{s}v$. In this case, we may choose L to be the identity $L(x, t)v = v$.

Suppose that the r.v. Y_t and ξ_n all have the same law. Indeed, in Equation (1), we have

$$\begin{aligned} X_{j+1}^{\tau^N} &= X_j^{\tau^N} + D\Big(X_j^{\tau^N}, Y_{t_j}, t_j, (t_{j+1} - t_j)\Big) \\ &= X_j^{\tau^N} + \sqrt{(t_{j+1} - t_j)}Y_{t_j} = X_j^{\tau^N} + \frac{1}{\sqrt{N}}Y_{t_j}. \end{aligned} \tag{75}$$

Then, we interpolate using

$$\begin{aligned} \mathfrak{X}_t^{\tau^N} &= X_j^{\tau^N} + D\Big(X_j^{\tau^N}, Y_{t_j}, t_j, (t - t_j)\Big), \\ X_j^{\tau^N} &+ \sqrt{(t - t_j)}Y_{t_j} = X_j^{\tau^N} + \sqrt{\frac{s}{N}}Y_{t_j}. \end{aligned} \tag{76}$$

for $j/N \le t \le (j+1)/N$; in the last step, we used equation (72). Hence, the processes $(\mathfrak{X}_t^{\tau^N})_t$ and (W^N) are identically distributed.

So, we can derive a corollary of Theorem 3 as follows.

Corollary 5. *Let $H = \mathbb{R}^d$, let $I = [0, T]$, and let $(Y_t)_{t \in \mathbb{R}^+}$ be i.i.d. H-valued random variables. We assume that they are centered and that, for the constants $c_t, c_3 > 0$ fixed in Hypotheses 1,*

$$\mathbb{E}\big[\|Y_t\|_H^4 e^{4c_3c_t\|Y_t\|_H}\big] < \infty. \tag{77}$$

Let $\mathfrak{X}_0$ be an H-valued random variable, independent of all Y_t, and we consider the processes $\mathfrak{X}^\tau$ defined as in Definition 9. We assume that the following clauses hold.

1.

$$\mathbb{E}\big[\|\mathfrak{X}_0\|_H^4\big] < \infty. \tag{78}$$

2. *There is a constant $c > 0$ such that for all $x, v \in H$, $t \in I$, and $s \in [0, 1]$,*

$$\|D(x, v, t, s) - \sqrt{s}\,v\|_H \le cs\big(\|x\|_H + 1\big)\|v\|_H^2. \tag{79}$$

3. *There is a constant $c_d > 1$ such that for all $x, v \in H$, $t \in \mathbb{R}^+$, and $s \in [0, 1]$,*

$$\begin{aligned} &\|D(x, v, t, s)\|_H \le c_d\sqrt{s}\|v\|_H, \\ &\sqrt{s}\|v\|_H \le 1/c_d \Longrightarrow \|D(x, v, t, s) - \sqrt{s}\,v\|_H \le c_d s\|v\|_H^2. \end{aligned} \tag{80}$$

Then, the family $\mathfrak{X}^\tau$, for $\tau \in \mathfrak{T}_T$, is tight in $C(I; H)$.

Proof 3. Since H is finite-dimensional, the identity $K = \mathrm{Id}_H$ is a compact operator and $K(H) = H$. We choose $L(x, t) = \mathrm{Id}_H$. Then, Conditions (2), (2L), and (3L) in Theorem 3 are automatic; the remaining assumptions are exactly those of Theorem 3. Hence, Theorem 3 applies and yields the conclusion. □

In the special case $D(x, v, t, s) = \sqrt{s}v$, we recover the tightness part of Donsker's theorem. The corollary, therefore, extends that tightness result to a rescaled nonlinear random walk. All the corollaries stated in Subsection 5.2.1 then apply.

We add an important remark.

Proposition 5. *There is no choice of common probability space where to define Y_t and W_t and such that the approximating processes W^N defined above would converge to W in probability.*

5.4 | Proof of Theorem 3

The proof is organized as follows. The first step establishes the compactness requirement (first hypothesis in Theorem 2) at fixed times, using the stronger $K(H)$-topology: Lemma 2 gives recursive K-moment bounds for the discrete variables X_n^τ, and Corollary 6 transfers these estimates to the interpolated process $\mathfrak{X}^\tau$, yielding the compact sets required in the first clause of Theorem 2. The second step controls the modulus of continuity in the ambient Hilbert norm. We first prove second- and fourth-moment bounds for increments (Lemmas 3 and 4) and then combine them with an Etemadi-type maximal estimate adapted to the random walk and to its interpolation (Lemma 6 and Corollary 7) to eventually obtain the probability bound in Equation (42). The following three subsections are organized according to this scheme.

5.4.1 | Proof of Theorem 3: Step 1

In this section, we prove that, under the hypotheses of Theorem 3, the first hypothesis in Theorem 2 is satisfied. We will use this Lemma in two ways, with K being the compact operator defined above in the hypotheses of Theorem 3, but also with K being the identity.

Lemma 2. *Let $K\colon H \longrightarrow H$ a linear injective operator (not necessarily compact). Let $\|v\|_K$ be defined in equation (50). We assume that for all $x, v \in K(H), t, s > 0$,*

$$D(x, v, t, s) \in K(H), L(x, t)v \in K(H), \tag{81}$$

(these are items 2 and 3 of Theorem 3), while we rewrite item 5 in this form: There are constants $c_{1D}, c_{2D} \geq 0$ such that for all $x, v \in K(H)$, $s \in [0, 1]$,

$$\begin{aligned}&\|D(x, v, t, s) - sL(x, t)v\|_K \\ &\quad\leq s\big(c_{1D}\|x\|_K\|v\|_H^2 + c_{2D}\|v\|_K\|v\|_H\big)e^{c_3\sqrt{s}\|v\|_H}.\end{aligned} \tag{82}$$

and we rewrite item 4: There are constants $c_{1L}, c_{2L} \geq 0$ such that for all $x, v \in K(H)$,

$$\|L(x, t)v\|_K \leq \big(c_{1L}\|x\|_K\|v\|_H + c_{2L}\|v\|_K\big)e^{c_3\|v\|_H}. \tag{83}$$

We define the following objects: let $\tau \in \mathfrak{T}$; fix $m \geq 0$ and 10 $F \in \mathcal{F}_{t_m}$ with $\mathbb{P}(F) > 0$; we will write $\mathbb{E}_F$ for the expectation computed using the conditional probability $\mathbb{P}(\cdot|F)$; we consider $n \geq m$; let

$$e_m = \mathbb{E}_F\Big[\big\|X_m^\tau\big\|_K^2\Big], b_n \stackrel{def}{=} \mathbb{E}_F\Big[\big\|X_n^\tau - X_m^\tau\big\|_K^2\Big]. \tag{84}$$

Then, we have two theses.

- *If $c_{1D} = c_{1L} = 0$, then*

$$b_n \leq \big(e^{c_5(t_n - t_m)} - 1\big). \tag{85}$$

- *Instead if $(c_{1L} + c_{1D}) > 0$, then*

$$b_n \leq (e_m + 1)\big(e^{c_5(t_n - t_m)} - 1\big). \tag{86}$$

where $c_5 > 0$ depends only on $c_{1L}, c_{2L}, c_{1D}, c_{2D}$, on $\widetilde{c}(4, 4c_3)$ from Proposition 1, on K and the law γ of Y_t, but c_5 does not depend on e_m, on F, and on τ.

Remark 10. Recall that

$$\sqrt{\mathbb{E}_F\Big[\big\|X_n^\tau\big\|_K^2\Big]} \leq \sqrt{\mathbb{E}_F\Big[\big\|X_n^\tau - X_m^\tau\big\|_K^2\Big]} + \sqrt{\mathbb{E}_F\Big[\big\|X_m^\tau\big\|_K^2\Big]}, \tag{87}$$

(or otherwise using Lemma 12 in Appendix A), we get

$$\mathbb{E}_F\Big[\big\|X_n^\tau\big\|_K^2\Big] \leq \Big(\sqrt{e_m} + \sqrt{b_n}\Big)^2 \leq 2(e_m + b_n). \tag{88}$$

Proof of Lemma 2: For $t \geq t_m$, we define

$$a_{K,\alpha} = \mathbb{E}_F\big[\|Y_t\|_K^\alpha\big] = \mathbb{E}\big[\|Y_t\|_K^\alpha\big], a_\alpha = \mathbb{E}_F\big[\|Y_t\|^\alpha\big] = \mathbb{E}\big[\|Y_t\|^\alpha\big]. \tag{89}$$

where the equality derives by independence; by Hypothesis 1, these are finite for $\alpha \leq 4$. For readability, we write X_n instead of X_n^τ; we write $\widetilde{X}_n$ instead of $X_n^\tau - X_0^\tau$ and $\delta_n = t_{n+1} - t_n$; we abbreviate

$$D_n = D\Big(X_n, Y_{t_n}, t_n, \sqrt{\delta_n}\Big), \tag{90}$$

$$A_n = D_n - \sqrt{\delta_n}L(t_n, X_n)Y_{t_n}. \tag{91}$$

By Equations (88) and (82), using Lemma 12 and Proposition 1, when $c_3 > 0$,

$$\begin{aligned}\mathbb{E}_F\big[\|A_n\|_K^2\big] &\leq 2c_{1D}^2\delta_n\mathbb{E}_F\Big[\|X_n\|_K^2\big\|Y_{t_n}\big\|_H^4 e^{2c_3\sqrt{\delta_n}\|Y_{t_n}\|_H}\Big] \\ &\quad + 2c_{2D}^2\delta_n\mathbb{E}_F\Big[\big\|Y_{t_n}\big\|_K^2\big\|Y_{t_n}\big\|_H^2 e^{2c_3\sqrt{\delta_n}\|Y_{t_n}\|_H}\Big] \\ &\leq 2c_{1D}^2\mathbb{E}_F\big[\|X_n\|_K^2\big]\mathbb{E}\Big[\Big\|\sqrt{\delta_n}\,Y_{t_n}\Big\|_H^4 e^{2c_3\|\sqrt{\delta_n}\,Y_{t_n}\|_H}\Big] \\ &\quad + 2c_{2D}^2\delta_n\sqrt{\mathbb{E}\Big[\big\|Y_{t_n}\big\|_K^4\Big]}\sqrt{\mathbb{E}\Big[\Big\|\sqrt{\delta_n}\,Y_{t_n}\Big\|_H^4 e^{4c_3\|\sqrt{\delta_n}\,Y_{t_n}\|_H}\Big]} \\ &\leq \delta_n^2\Big(4c_{1D}^2(b_n + e_m)\widetilde{c}(4, 2c_3) + 2c_{2D}^2\sqrt{\widetilde{c}(4, 4c_3)a_{4,K}}\Big).\end{aligned} \tag{92}$$

Summarizing, we have

$$\mathbb{E}_F\big[\|A_n\|_K^2\big] \leq \delta_n^2 c_4^2\big(c_{1D}^2(b_n + e_m) + 1\big). \tag{93}$$

where c_4 depends only on c_{2D}, c_3, on $\widetilde{c}$ from Proposition 1, and on $c_t > 0$ from Hypothesis 1 and Definition 8, but c_4 does not depend on F, c_{1D}, and τ.

Similarly, using (83),

$$\begin{aligned}\delta_n\mathbb{E}_F\Big[\big\|L(X_n, t)Y_{t_n}\big\|_K^2\Big] &\leq 2c_{1L}^2\delta_n\mathbb{E}_F\Big[\|X_n\|_K^2\big\|Y_{t_n}\big\|_H^2 e^{2c_3\sqrt{\delta_n}\|Y_{t_n}\|_H}\Big] \\ &\quad + 2c_{2L}^2\delta_n\mathbb{E}\Big[\big\|Y_{t_n}\big\|_K^2 e^{2c_3\sqrt{\delta_n}\|Y_{t_n}\|_H}\Big] \\ &\leq 2c_{1L}^2\mathbb{E}_F\big[\|X_n\|_K^2\big]\mathbb{E}\Big[\Big\|\sqrt{\delta_n}\,Y_{t_n}\Big\|_H^2 e^{2c_3\|\sqrt{\delta_n}\,Y_{t_n}\|_H}\Big] \\ &\quad + 2c_{2L}^2\delta_n\sqrt{\mathbb{E}\Big[\big\|Y_{t_n}\big\|_K^4\Big]}\sqrt{\mathbb{E}\Big[e^{4c_3\|\sqrt{\delta_n}\,Y_{t_n}\|_H}\Big]} \\ &\leq \delta_n\Big(4c_{1L}^2(b_n + e_m)\widetilde{c}(2, 2c_3) + 2c_{2L}^2\sqrt{\widetilde{c}(0, 4c_3)a_{4,K}}\Big),\end{aligned} \tag{94}$$

summarized to

$$\delta_n\mathbb{E}_F\Big[\big\|L(X_n, t)Y_{t_n}\big\|_K^2\Big] \leq \delta_n c_4^2\big(c_{1L}^2(b_n + e_m) + 1\big), \tag{95}$$

possibly enlarging c_4, that now depends also on c_{2L}.

We estimate iteratively. We begin by expressing

$$
\begin{aligned}
\left\|\widetilde{X}_{n+1}\right\|_K^2 &= \left\|D_n+\widetilde{X}_n\right\|_K^2 = \left\|\widetilde{X_n}+\sqrt{\delta_n}\,L(X_n,t)Y_{t_n}+A_n\right\|_K^2 \\
&= \left\|\widetilde{X_n}\right\|_K^2 + \delta_n\left\|L(X_n,t)Y_{t_n}\right\|_K^2 + \|A_n\|_K^2 \\
&\quad + 2\sqrt{\delta_n}\left\langle\widetilde{X_n}, L(X_n,t)Y_{t_n}\right\rangle_K + 2\left\langle\widetilde{X_n}, A_n\right\rangle_K \\
&\quad + 2\sqrt{\delta_n}\langle A_n, L(X_n,t)Y_{t_n}\rangle_K,
\end{aligned} \tag{96}
$$

and we then compute the expectation; we note that

$$
\mathbb{E}_F\left[\left\langle\widetilde{X_n}, L(X_n,t)Y_{t_n}\right\rangle_K\right] = 0, \tag{97}
$$

because Y_{t_n} has zero average and is independent of F, X_n, and X_m, whereas

$$
\begin{aligned}
\mathbb{E}_F\left[\left\|\widetilde{X}_n\right\|_K^2\right] &= b_n, \\
\delta_n\mathbb{E}_F\left[\left\|L(X_n,t)Y_{t_n}\right\|_K^2\right] &\le \delta_n c_4^2\left(c_{1L}^2(b_n+e_m)+1\right), \\
\mathbb{E}_F[\|A_n\|_K^2] &\le \delta_n^2 c_4^2\left(c_{1D}^2(b_n+e_m)+1\right), \\
\mathbb{E}_F\left[\left\langle A_n, \sqrt{\delta}_n L(X_n,t)Y_{t_n}\right\rangle_K\right] &\le \sqrt{\mathbb{E}_F[\|A_n\|_K^2]}\sqrt{\mathbb{E}_F\left[\delta_n\left\|L(X_n,t)Y_{t_n}\right\|_K^2\right]} \\
&\le \delta_n^{3/2}c_4^2\left(c_1^2(b_n+e_m)+1\right), \\
\mathbb{E}_F\left[\left\langle\widetilde{X}_n, A_n\right\rangle_K\right] &\le \sqrt{b_n}\sqrt{\mathbb{E}_F[\|A_n\|_K^2]},
\end{aligned} \tag{98}
$$

where $c_1 = \max\{c_{1L}, c_{1D}\}$.

In this last line, if $c_{1D} = c_{1L} = 0$, then again, we use $\sqrt{s} \le s+1$ and Equation (93), so

$$
\mathbb{E}_F\left[\left\langle\widetilde{X}_n, A_n\right\rangle_K\right] \le \sqrt{b_n}\sqrt{\mathbb{E}_F[\|A_n\|_K^2]} \le \delta_n c_4\sqrt{b_n} \le \delta_n c_4(b_n+1), \tag{99}
$$

so (recalling that $\delta_n \le c_t$, the constant from Hypothesis 1), we estimate as follows:

$$
b_{n+1} \le b_n + c_5\delta_n(1+b_n) = b_n(1+c_5\delta_n) + c_5\delta_n, \tag{100}
$$

which, by Lemma 15 in Appendix A (shifting the sequence), implies Equation (85).

Instead, if $c_1 > 0$, we note that

$$
s\left(c_1^2 s + 2a\right) \le c_1^2 s^2 + 2as + \frac{a^2}{c_1^2} = \left(c_1 s + \frac{a}{c_1}\right)^2, \tag{101}
$$

hence using equation (93) and setting $s = b_n$, $a = (c_1^2 e_m + 1)/2$,

$$
\begin{aligned}
\mathbb{E}_F\left[\left\langle\widetilde{X}_n, A_n\right\rangle_K\right] &\le \delta_n c_4\sqrt{b_n}\sqrt{\left(c_1^2(b_n+e_m)+1\right)} \\
&\le \delta_n c_4\left(c_1 b_n + c_1\frac{e_m}{2} + \frac{1}{2c_1}\right).
\end{aligned} \tag{102}
$$

Eventually, we estimate as follows:

$$
b_{n+1} \le b_n(1+c_5\delta_n) + c_5\delta_n(e_m+1), \tag{103}
$$

which, by Lemma A.4 in Appendix A (shifting the sequence), implies equation (86).

Corollary 6. *Let $m = 0$, $F = \Omega$, then by equations (88) and (86), we obtain that*

$$
\mathbb{E}\left[\left\|X_n^\tau\right\|_K^2\right] \le (e_0+1)e^{c_5 t_n}, \tag{104}
$$

with $e_0 = \mathbb{E}[\|\mathfrak{X}_0\|_K^2] < \infty$, under assumption (52); moreover, as explained in Remark 6, we have

$$
\mathbb{E}\left[\left\|\mathfrak{X}_t^\tau\right\|_K^2\right] \le (e_0+1)e^{c_5 t}, \tag{105}
$$

for all $t \ge 0$.

Conclusion of Step 1: So, to conclude the first step, we consider a process $\mathfrak{X}^\tau$ and a time $T > 0$; let $\varepsilon > 0$, for $0 \le t \le T$, $r > 0$, we have then by Markov inequality,

$$
\mathbb{P}\left\{\left\|\mathfrak{X}_t^\tau\right\|_K \ge r\right\} \le \frac{\mathbb{E}\left[\left\|\mathfrak{X}_t^\tau\right\|_K^2\right]}{r^2} \le \frac{1}{r^2}c_K, \tag{106}
$$

with

$$
c_K = (e_0+1)e^{c_5 T}, \tag{107}
$$

so setting $r_j = \sqrt{(c_K 2^j/\varepsilon)}$,

$$
\mathbb{P}\left\{\mathfrak{X}_{a_j}^\tau \notin D_K(0, r_j)\right\} < \varepsilon 2^{-j}, \tag{108}
$$

and this satisfies the first hypothesis in Theorem 2.

5.4.2 | Proof: Lemmas for Step 2

In this section, we prove some powerful lemmas that then will be used to prove that the second hypothesis in Theorem 2 is satisfied.

Remark 11. We consider item 6 of Theorem 3; note that, for $s \in [0,1]$, by Equations (58) and (60),

$$
\begin{aligned}
\|D(x,v,t,s) - \sqrt{s}\,L(x,t)v\|_H &\le \|D(x,v,t,s)\|_H + \|\sqrt{s}\,L(x,t)v\|_H \\
&\le 2c_d\sqrt{s}\|v\|_H \le 2c_d^2 s\|v\|_H^2,
\end{aligned} \tag{109}
$$

for $\sqrt{s}\|v\| \ge 1/c_d$, so adding (59), we obtain

$$
\forall x, v \in H, \forall s \in [0,1], t \ge 0, \|D(x,v,t,s) - \sqrt{s}\,L(x,t)v\|_H \le 2c_d^2 s\|v\|_H^2. \tag{110}
$$

Lemma 3. *We assume item 6 of Theorem 3, define the objects as in Lemma 2, with K being the identity, and recall that in this case,*

$$
b_n \overset{def}{=} \mathbb{E}_F\left[\left\|X_n^\tau - X_m^\tau\right\|_H^2\right]; \tag{111}
$$

then, for $n \ge m$,

$$
b_n \le \left(e^{c_5(t_n - t_m)} - 1\right). \tag{112}
$$

In particular, for $0 \le t_m \le t_n \le T$, we have

$$
b_n \le c_6(t_n - t_m). \tag{113}
$$

Proof 4. We use Remark 11 and Remark 6 and apply Lemma 2 with K being the identity, $c_{1L} = c_{1D} = c_3 = 0$ and $c_{2L} = c_{2D} = 2c_d^2$; we obtain the constant $c_5 > 0$, and hence, we set $c_6 = c_5 e^{c_5 T}$. ☐

Lemma 4. *We assume item 6 of Theorem 3; we fix $\tau \in \mathfrak{T}$; we fix $m \ge 0$ and $F \in \mathcal{F}_{t_m}$; we will write $\mathbb{E}_F$ for the expectation computed* *using the conditional probability $\mathbb{P}(\cdot|F)$; we consider $\mathbb{P}(\cdot|F)$; and letting*

$$b_n \overset{def}{=} \mathbb{E}_F\left[\left\| X_n^\tau - X_m^\tau \right\|_H^2 \right], \tag{114}$$

suppose that there is a constant c_6 such that

$$b_n \le c_6(t_n - t_m). \tag{115}$$

for $0 \le t_m \le t_n \le T$ (as in equation (113)) and eventually let

$$q_n \overset{def}{=} \mathbb{E}_F\left[\left\| X_n^\tau - X_m^\tau \right\|_H^4 \right]. \tag{116}$$

We prove that, for $0 \le t_n \le t_m \le T$,

$$q_n \le (c_7 + 2c_8)g(t_n - t_m), \tag{117}$$

where

$$g(t) = \frac{e^{c_7 t} - 1 - c_7 t}{c_7^2}, \tag{118}$$

and where c_8, c_7 depend only on c_6, c_d and the fourth moment of Y_t.

Proof of Lemma: For $t \ge t_m$, we define

$$a_\alpha = \mathbb{E}_F[\|Y_t\|^\alpha] = \mathbb{E}[\|Y_t\|^\alpha], \tag{119}$$

where the equality derives by independence.

Again, for readability, we write X_n instead of X_n^τ, $\widetilde{X}_n$ instead of $\widetilde{X}_n$, $\delta_n = t_{n+1} - t_n$, and

$$D_n = D\left(X_n, Y_{t_n}, t_n, \sqrt{\delta_n}\right), \tag{120}$$

and $\|x\| = \|x\|_H$. We compute

$$\begin{aligned} \left\|\widetilde{X}_{n+1}\right\|^4 &= \left\|\widetilde{X}_n + D_n\right\|^4 = \left(\left\|\widetilde{X}_n\right\|^2 + 2\left\langle\widetilde{X}_n, D_n\right\rangle + \|D_n\|^2\right)^2 \\ &= \left\|\widetilde{X}_n\right\|^4 + 4\left\langle\widetilde{X}_n, D_n\right\rangle^2 + \|D_n\|^4 \\ &\quad + 4\left\|\widetilde{X}_n\right\|^2\left\langle\widetilde{X}_n, D_n\right\rangle + 4\left\langle\widetilde{X}_n, D_n\right\rangle\|D_n\|^2 \\ &\quad + 2\left\|\widetilde{X}_n\right\|^2\|D_n\|^2, \end{aligned} \tag{121}$$

then we compute integrals; for the fourth term, since

$$\mathbb{E}_F\left[\left\|\widetilde{X}_n\right\|^2\left\langle\widetilde{X}_n, L(X_n, t_n)Y_{t_n}\right\rangle\right] = 0, \tag{122}$$

by independence, then

$$\mathbb{E}_F\left[\left\|\widetilde{X}_n\right\|^2\left\langle\widetilde{X}_n, D_n\right\rangle\right] = \mathbb{E}_F\left[\left\|\widetilde{X}_n\right\|^2\left\langle\widetilde{X}_n, D_n - \sqrt{\delta_n}\, L(X_n, t_n)Y_{t_n}\right\rangle\right], \tag{123}$$

so by Equation (110),

$$\begin{aligned} \left|\mathbb{E}_F\left[\left\|\widetilde{X}_n\right\|^2\left\langle\widetilde{X}_n, D_n\right\rangle\right]\right| &\le \mathbb{E}_F\left[\left\|\widetilde{X}_n\right\|^3\left\|D_n - \sqrt{\delta_n}\, L(X_n, t_n)Y_{t_n}\right\|\right] \\ &\le 2c_d^2\delta_n\mathbb{E}_F\left[\left\|\widetilde{X}_n\right\|^3\left\|Y_{t_n}\right\|^2\right] \\ &= 2c_d^2\delta_n\mathbb{E}_F\left[\left\|\widetilde{X}_n\right\|^3\right]\mathbb{E}\left[\left\|Y_{t_n}\right\|^2\right] \\ &\le 2c_d^2\delta_n a_2\sqrt{\mathbb{E}_F\left[\left\|\widetilde{X}_n\right\|^2\right]\mathbb{E}_F\left[\left\|\widetilde{X}_n\right\|^4\right]} \\ &\le 2c_d^2\delta_n a_2\sqrt{q_n}\sqrt{b_n}, \end{aligned} \tag{124}$$

(again by using independence in the third step). For the other terms, using Equation (58),

$$\begin{aligned} \mathbb{E}_F\left[\left\langle\widetilde{X}_n, D_n\right\rangle^2\right] &\le \mathbb{E}_F\left[\left\|\widetilde{X}_n\right\|^2\|D_n\|^2\right] \\ &\le c_d^2\mathbb{E}_F\left[\left\|\widetilde{X}_n\right\|^2\left\|\sqrt{\delta_n}\, Y_{t_n}\right\|^2\right] \le c_d^2 a_2 b_n\delta_n \\ \mathbb{E}_F\left[\|D_n\|^4\right] &\le c_d^4 a_4\delta_n^2, \\ \mathbb{E}_F\left[\left\langle\widetilde{X}_n, D_n\right\rangle\|D_n\|^2\right] &\le \mathbb{E}_F\left[\left\|\widetilde{X}_n\|\|D_n\right\|^3\right] \\ &\le c_d^3\delta_n^{3/2}\mathbb{E}_F\left[\left\|\widetilde{X}_n\|\|Y_n\right\|^3\right] \le c_d^3 a_3\sqrt{b_n}\delta_n^{3/2}. \end{aligned} \tag{125}$$

Eventually, we use $b_n \le c_6 t_n$ and note that $t_n \ge \delta_n$ then $t_n\delta_n \ge \sqrt{t_n}\delta_n^{3/2}$ and $t_n\delta_n \ge \delta_n^2$; hence (defining $c_7, c_8 > 0$ appropriately), summarizing

$$q_{n+1} \le q_n + \left(c_7\sqrt{q_n t_n} + c_8 t_n\right)\delta_n, \tag{126}$$

using Lemma 14 in Appendix A (shifting the sequence), we obtain equation (117).

We recall Etemadi's inequality [19] in the version of Theorem 22.5 in [20].

Lemma 5 (Etemadi's inequality). *Suppose that S_n is a process taking value in normed space, and it is the sum $S_n = Y_1 + \cdots + Y_n$ of i.i.d. r.v. $(Y_n)_n$; then, for $\varepsilon > 0$, we have*

$$\mathbb{P}\left(\max_{1\le k\le l}|S_k| \ge 3\varepsilon\right) \le 3\max_{1\le k\le l}\mathbb{P}(|S_k| \ge \varepsilon). \tag{127}$$

This is a key step in the proof of Donsker theorem, but we cannot use it in this form. To conclude the proof of Theorem 3, we need to prove a similar result, adapted to our process and hypotheses.

Lemma 6. *In the hypotheses of Theorem 3, with $\tau \in \mathfrak{T}$ (as in previous Lemmas), we fix $T, \varepsilon > 0$, and we fix $l \ge m > 0$ integers such that $t_l \le T$, then*

$$\mathbb{P}\left(\max_{m\le k\le l}\left\|X_k^\tau - X_{m-1}^\tau\right\| > 3\varepsilon\right) \le \frac{c_{10}}{\varepsilon^4}g(t_l - t_{m-1}), \tag{128}$$

where again g was defined in Equation (118), and c_{10} depends only on the constants in previous lemmas.

Proof of Lemma 5. Let $X = X^\tau$ for simplicity, and

$$\widetilde{X}_k = X_k - X_{m-1}, \tag{129}$$

and note that

$$\widetilde{X}_n - \widetilde{X}_j = X_n - X_j. \tag{130}$$

Let

$$A_m = \left\{\left\|\widetilde{X}_m\right\| > 3\varepsilon\right\}, \tag{131}$$

and for $j = m+1, \ldots l$, let

$$A_j = \left\{\max_{m\le i\le j-1}\left\|\widetilde{X}_i\right\| \le 3\varepsilon \wedge \left\|\widetilde{X}_j\right\| > 3\varepsilon\right\}, \tag{132}$$

so

$$\bigcup_{j=m}^{l} A_j = \left\{ \max_{m\le k\le l} \left\| \widetilde{X}_k \right\| > 3\varepsilon \right\}, \tag{133}$$

then further disintegrating,

$$\begin{aligned}\mathbb{P}\left(\max_{m\le k\le l}\left\|\widetilde{X}_k\right\| > 3\varepsilon\right) &\le \mathbb{P}\left(\left\|\widetilde{X}_l\right\| \ge \varepsilon\right) + \sum_{j=m}^{l}\mathbb{P}\left(A_j \cap \left\{\left\|\widetilde{X}_l\right\| < \varepsilon\right\}\right)\\ &\le \mathbb{P}\left(\left\|\widetilde{X}_l\right\| \ge \varepsilon\right) + \sum_{j=m}^{l}\mathbb{P}\left(A_j \cap \left\{\left\|X_l - X_j\right\| > 2\varepsilon\right\}\right)\\ &= \mathbb{P}\left(\left\|\widetilde{X}_l\right\| \ge \varepsilon\right) + \sum_{j=1}^{l}\mathbb{P}(A_j)\mathbb{P}\left(\left\{\left\|X_l - X_j\right\| > 2\varepsilon\right\}\middle|A_j\right),\end{aligned} \tag{134}$$

by Markov,

$$\mathbb{P}\left(\left\{\left\|X_l - X_j\right\| > 2\varepsilon\right\}\middle|A_j\right) \le \frac{1}{2^4\varepsilon^4}\mathbb{E}\left[\left\|X_l - X_j\right\|^4\middle|A_j\right]. \tag{135}$$

We use Lemma 3 with K identity and $F = A_j$; note that indeed $A_j \in \mathcal{F}_{t_j}$; having set $T > 0$, we obtain equation (112), that is, equation (115). So, Equation (115) is satisfied, and we can apply Lemma 4 to obtain Equation (117) in the thesis in Lemma 4 that we rewrite as

$$\mathbb{E}\left[\left\|X_n - X_j\right\|^4 | A_j\right] \le (c_7 + 2c_8)g\left(t_n - t_j\right). \tag{136}$$

Plugging it all in

$$\sum_{j=m}^{l}\mathbb{P}(A_j)\mathbb{P}\left(\left\{\left\|X_l - X_j\right\| > 2\varepsilon\right\}\middle|A_j\right) \le \frac{1}{2^4\varepsilon^4}(c_7 + 2c_8)g(t_l - t_{m-1})\sum_{j=m}^{l}\mathbb{P}(A_j). \tag{137}$$

Similarly, we deal with the first term $\mathbb{P}(\|\widetilde{X}_l\| \ge \varepsilon)$. □

We then prove the same result for the process $\mathfrak{X}^\tau$.

Corollary 7. *In the hypotheses of the previous lemma, we fix $\varepsilon > 0$, then for $t_m \in \tau$ and $t_m \le s \le T$,*

$$\mathbb{P}\left(\sup_{t_m<t\le s}\left\|\mathfrak{X}^\tau_t - \mathfrak{X}^\tau_{t_m}\right\| > 3\varepsilon\right) \le \frac{c_{10}}{\varepsilon^4}g(s - t_m). \tag{138}$$

Proof 5. We fix $\tau \in \mathfrak{T}_T$ and $t_m \in \tau$. For $t_m \le s \le T$, let

$$A^\tau_s = A_s \stackrel{\text{def}}{=} \left\{\sup_{t_m\le t\le s}\left\|\mathfrak{X}^\tau_t - \mathfrak{X}^\tau_{t_m}\right\| > 3\varepsilon\right\}, \tag{139}$$

then

$$A_s = \left\{\sup_{t_m\le t\le s, t\in\mathbb{Q}}\left\|\mathfrak{X}^\tau_t - \mathfrak{X}^\tau_{t_m}\right\| > 3\varepsilon\right\}, \tag{140}$$

since trajectories are continuous. Then, for $s_1 < s_2$, we have $A_{s_1} \subseteq A_{s_2}$, and moreover,

$$\bigcup_{s_1<s_2} A_{s_1} = A_{s_2}, \tag{141}$$

again using the fact that trajectories are continuous; hence, we obtain left-continuity

$$\sup_{s_1<s_2}\mathbb{P}\left(A_{s_1}\right) = \lim_{s_1\longrightarrow s_2-}\mathbb{P}\left(A_{s_1}\right) = \mathbb{P}\left(A_{s_2}\right). \tag{142}$$

□

As noted in Remark 1, if $\widehat{t} > 0$ (and not necessarily $\widehat{t} \in \mathbb{Q}$), if we add $\widehat{t}$ to τ and obtain $\widehat{\tau} = \tau \cup \{\widehat{t}\}$, then

$$\mathfrak{X}^{\widehat{\tau}}_t = \mathfrak{X}^\tau_t \quad \forall t \le \widehat{t}, \tag{143}$$

so

$$A^{\widehat{\tau}}_t = A^\tau_t \quad \forall t \le \widehat{t}, \tag{144}$$

but then we can apply the previous lemma and the above left-continuity to say that

$$\mathbb{P}\left(A^{\widehat{\tau}}_t\right) = \mathbb{P}\left(A^\tau_t\right) \le \frac{c_{10}}{\varepsilon^4}g(t - t_m). \tag{145}$$

We recall this other fundamental lemma (that is the key to Theorem 8.3 in [18]).

Lemma 7. *Suppose E is a normed vector space, $I = [a, b]$; let $\eta, \varepsilon, \nu > 0$ with $\nu \in \mathbb{N}$; suppose μ is a probability measure on the space $C = C(I; E)$, and let $a = s_0 < s_1 < \ldots s_\nu = b \in I$ with*

$$(s_{i+1} - s_i) \ge \eta, \quad \text{for } i = 2, \ldots \nu - 2, \tag{146}$$

then

$$\mu\{x \in C: \omega(x, \eta) \ge 3\varepsilon\} \le \sum_{i=0}^{\nu-1}\mu\left\{x \in C: \sup_{s_i\le s<s_{i+1}}\|x(s) - x(s_i)\|_E \ge \varepsilon\right\}. \tag{147}$$

For the proof, see the Corollary after Theorem 8.3 in [18]. Note that the inequality (146) need not hold for $i = 1, i = \nu - 1$.

5.4.3 | Proof of Theorem 3: Step 2

Now that we have proved the powerful lemmas, we prove the second hypothesis in Theorem 2, that is, Equation (42); to this end, we fix $\varepsilon_0 > 0, \varepsilon_1 > 0$; there is an $\eta > 0$ with $\eta \in \mathbb{Q}, \eta < c_t, \eta < T/2$ such that

$$9^4\ulcorner\frac{T}{\eta}\urcorner c_{10}\frac{g(\eta)}{\varepsilon_0^4} < \varepsilon_1, \tag{148}$$

where g was defined in Equation (118) in Lemma 4; Equation (42) will be satisfied with $A = \mathfrak{T}_T$ and

$$\alpha_0 = \tau_0 = \{\eta i: 0 \le i < \nu\} \cup \{T\}. \tag{149}$$

where $\nu = \lceil T/\eta \rceil$.

Let $\varepsilon = \varepsilon_0/9$. We define for convenience $s_i = \eta i$ (that are equispaced for $i < \nu$) while $s_\nu = T$. We consider any $\tau \supseteq \tau_0$; let $\mathfrak{X}^\tau$ be a process; by Equation (147),

$$\mathbb{P}\{\omega(\mathfrak{X}^\tau, \eta) \ge 9\varepsilon\} \le \sum_{i=0}^{\nu-1}\mathbb{P}\left\{\sup_{s_i\le t<s_{i+1}}\left\|\mathfrak{X}^\tau_t - \mathfrak{X}^\tau_{s_i}\right\|_H \ge 3\varepsilon\right\}. \tag{150}$$

For the terms in the sum in Equation (150), we use our version (Equation (138)) of Etemadi's estimate to obtain

$$\begin{aligned}\mathbb{P}\{\omega(\mathfrak{X}^\tau, \eta) \ge 9\varepsilon\} &\le \sum_{i=0}^{\nu-1}\mathbb{P}\left\{\sup_{s_i\le t<s_{i+1}}\left\|\mathfrak{X}^\tau_t - \mathfrak{X}^\tau_{s_i}\right\|_H > 3\varepsilon\right\}\\ &\le \nu\frac{c_{10}}{\varepsilon^4}g(\eta) < \varepsilon_1\end{aligned} \tag{151}$$

by Equation (148). So, we have satisfied the second hypothesis of Theorem 2, in the form expressed in Equation (43).

This concludes the proof of Theorem 3.

6 | Results on Manifolds

6.1 | Hypotheses for Manifolds

We again define, for $v \in K(H)$, $\|v\|_K \overset{def}{=} \|K^{-1}v\|_H$ as in Definition 11; similarly, for scalar products. The following theorem uses the following hypotheses on the manifold M and its embedding in H. Let $\exp_x v$ be the exponential mapping of M. For convenience, we denote by $P_x: H \longrightarrow T_xM$ the orthogonal projection $P_x = \pi_{T_xM}$.

Hypotheses 2. We assume what follows.

1. The manifold M is isometrically embedded in the Hilbert space H, and it is a closed subset of it.
2. The second fundamental form of the embedding of the manifold M is uniformly bounded.

We suppose that there exists a compact operator K satisfying the requisites in the previous Section 5.1, and constants $c_e, c_p > 0$, such that

3.

$$\mathbb{P}\{\mathfrak{X}_0 \in M\} = 1, \tag{152}$$

and

$$\mathbb{E}\left[\|\mathfrak{X}_0\|_K^4\right] < \infty. \tag{153}$$

4. If $x \in M \cap K(H)$ and $v \in K(H)$, then $P_x v \in K(H)$
5. and

$$\|P_x v\|_K \le c_p\big(\|v\|_K + \|x\|_K\|v\|_H\big); \tag{154}$$

6. If $x \in M \cap K(H)$ and $v \in T_xM \cap K(H)$, then $\exp_x v \in M \cap K(H)$
7. and

$$\big\|(\exp_x v) - (v + x)\big\|_K \le c_e\big(\|x\|_K\|v\|_H^2 + \|v\|_K\|v\|_H\big)e^{c_3\|v\|_H}. \tag{155}$$

The second hypothesis can be reformulated as follows.

Proposition 6. *The following are equivalent:*

1. *The second fundamental form of the embedding of manifold M is uniformly bounded.*
2. *$\exists c_e > 0$ such that $\forall x \in M, \forall v \in T_xM$,*

$$\|v\|_H \le \frac{1}{c_e} \Longrightarrow \big\|(\exp_x v) - (v + x)\big\|_H \le c_e\|v\|_H^2. \tag{156}$$

6.2 | Tightness of Interpolated Geodesic Processes

Theorem 5. *We consider the interpolated geodesic processes $\mathfrak{X}^\tau$ defined as in Section 3.5.1; we restrict each $\mathfrak{X}^\tau_t$ to $t \in I = [0, T]$. Under Hypothesis 2, these $\mathfrak{X}^\tau$, for $\tau \in \mathfrak{T}_T$, are a tight family in $C(I; M)$.*

Since this theorem is proved using Theorem 3, then all corollaries of the latter hold also for the former. We have, moreover, this result.

Corollary 8. *Any limit point $\mathfrak{X}$ of $\mathfrak{X}^\tau$ is a process taking values in $C(I; M)$ a.s.*

Proof 6. Since $M \subset H$ was assumed to be closed, then $C(I; M)$ is a closed subset of $C(I; H)$; by construction, $\mathfrak{X}^\tau_t \in M$ for all $t \in I$, hence

$$\mathbb{P}(\mathfrak{X}^\tau \in C(I; M)) = 1, \tag{157}$$

so by Alexandrov's theorem[11],

$$\mathbb{P}(\mathfrak{X} \in C(I; M)) = 1. \tag{158}$$

□

Remark 12. Nothing is specifically "infinite dimensional" in this approach: This theorem can be applied to finite dimensional manifolds as well. Recall that, by Nash embedding theorems, any finite dimensional Riemannian manifold can be isometrically embedded in $H = \mathbb{R}^N$, and in this case, we set K to be the identity; moreover, Equation (154) is trivially true. We then require that

$$\mathbb{P}\{\mathfrak{X}_0 \in M\} = 1, \mathbb{E}\left[|\mathfrak{X}_0|^4\right] < \infty; \tag{159}$$

then, we require the bound on the second fundamental form that implies Equation (158) that in turn implies Equation (157) (see Remark 11); under these conditions, Theorem 5 holds.

Proof of Theorem 5: As in Equation (38) in Section 3.5.1, we define

$$D(x, y, t, s) \overset{\text{def}}{=} \exp_x\left(\sqrt{s}\, P_x y\right) - x, \tag{160}$$

and we define

$$L(x, t)v = P_x v. \tag{161}$$

In this way, the random walk on the manifold can be seen as a special case of a random walk in H. We recall item 6 of Theorem 3:

$$\|D(x, v, t, s)\|_H \le c_d\sqrt{s}\|v\|_H, \tag{162}$$

$$\sqrt{s}\|v\| \le \frac{1}{c_d} \Longrightarrow \|D(x, v, t, s) - \sqrt{s}\, L(x, t)v\|_H \le c_d s\|v\|_H^2. \tag{163}$$

The first one, when $c_d \ge 1$, is true for any embedded manifold since the length of a curve is larger (or equal) than the distance between its end points, and $\|P_x v\|_H \le \|v\|_H$. For the second one, for $\sqrt{s}\|v\| \le 1/c_e$, we write

$$\|D(x,v,t,s) - \sqrt{s}\,L(x,t)v\|_H = \left\| \exp_x\!\left(\sqrt{s}\,P_x v\right) - \left(\sqrt{s}\,P_x v + x\right) \right\|_H \leq c_e s \|P_x v\|_H^2 \leq c_e s \|v\|_H^2, \tag{164}$$

using (156) from Proposition 6.

To obtain item 5 of Theorem 3: we substitute in (155) to obtain

$$\left\| (\exp_x P_x v) - (P_x v + x) \right\|_K \leq c_e\left(\|x\|_K \|P_x v\|_H^2 + \|P_x v\|_K \|P_x v\|_H\right) e^{c_3 \|P_x v\|_H}. \tag{165}$$

Then, we use $\|P_x v\|_H \leq \|v\|_H$ again, and we use Equation (154), so

$$\begin{aligned} &\left\| \exp_x(P_x v) - (P_x v + x) \right\|_K \\ &\quad \leq c_e \|x\|_K \|v\|_H^2 e^{c_3 \|v\|_H} + c_e c_p\left(\|v\|_K + \|x\|_K \|v\|_H\right)\|v\|_H e^{c_3 \|v\|_H} \\ &\quad = \left(c_e(1+c_p)\|x\|_K \|v\|_H^2 + c_e c_p \|v\|_K \|v\|_H\right) e^{c_3 \|v\|_H}. \end{aligned} \tag{166}$$

Then, replacing $\sqrt{s}v$ for v, this last satisfies Equation (57) with $c = c_e(1+c_p)$. So, Theorem 5 can be straightforwardly seen as a corollary of Theorem 3.

7 | Results on Stiefel Manifolds

Let V be a separable Hilbert space with $\dim(V) \geq 2p$ (possibly infinite dimensional). We define the Hilbert space V^p and the Stiefel manifold $\mathbf{St}(p,V)$ as in Section 2.2.

In the following two sections, we will show that the above Hypothesis 2 is satisfied when $M = \mathbf{St}(p,V)$ is the Stiefel manifold; so, the family of random walks is tight.

We will use these definitions with $E = H$ or $E = H^2$.

Definition 13. If E is a vector space with a scalar product, we agree that, for $x, v \in E^p$, $A = x^\top v$ is the $p \times p$ matrix defined by

$$A_{i,j} = \langle x_i, v_j \rangle_E. \tag{167}$$

We also agree that, given $x \in E^p$ and $A \in \mathbb{R}^{p\times p}$, the right product

$$y = xA \tag{168}$$

is the vector $y \in E^p$.

$$y_i = \sum_{j=1}^{p} x_j A_{j,i}. \tag{169}$$

7.1 | Probabilities on Stiefel Manifolds

When $M = \boldsymbol{St}(p,V)$ is a Stiefel manifold, it will be convenient to build the probabilistic infrastructure in Section 3.2 in this specific way.

Suppose that $\tilde{\gamma} = N(0,\tilde{Q})$ is a centered nondegenerate Gaussian measure in the separable Hilbert space V. We will then define the operator $Q\colon H \longrightarrow H$ by tensor product

$$\langle x, Qy \rangle_H = \sum_{i=1}^{p} \left\langle x_i, \tilde{Q} y_i \right\rangle_V, \tag{170}$$

so $\gamma = N(0,Q)$ is a centered nondegenerate Gaussian measure in the separable Hilbert space H, given by the measure product

$$\gamma = \tilde{\gamma} \otimes \cdots \otimes \tilde{\gamma}. \tag{171}$$

Equivalently, if we consider $x \in H$ as a r.v. with distribution γ, then the columns of $x \in H$ will be independent r.v. each with distribution $\tilde{\gamma}$.

Proposition 7. *Given $A \in O(p)$, the action*

$$A\colon H \longrightarrow H, \quad x \longmapsto xA \tag{172}$$

maps identically

$$\gamma = A_\sharp \gamma, \tag{173}$$

the probability γ to itself.

7.2 | Tightness Operator

As noted in Remark 4, in the space V starting from $\tilde{Q}$, we can define a compact operator $\tilde{K}\colon V \longrightarrow V$ such that $\tilde{K}^{-1}\tilde{Q}\tilde{K}^{-1}$ is still trace class. We eventually define $K\colon H \longrightarrow H$ as

$$y = Kx \text{when} y_i = \tilde{K} x_i. \tag{174}$$

So, K commutes with the right multiplication by matrices

$$(Kx)A = K(xA), \tag{175}$$

defined in Definition 13.

7.3 | Tightness in Stiefel Manifolds

Theorem 6. *Suppose that $Y_t \sim N(0,Q)$ as defined in previous Section 7.1. We 0consider interpolated geodesic processes $\mathfrak{X}^\tau$ defined as in Section 3.5.1 when M is a Stiefel manifold, having*

$$\mathbb{P}\{\mathfrak{X}_0 \in M\} = 1, \mathbb{E}\left[\|\mathfrak{X}_0\|_K^4\right] < \infty, \tag{176}$$

and restrict each $\mathfrak{X}_t^\tau$ to $t \in I = [0,T]$; these $\mathfrak{X}^\tau$, for $\tau \in \mathfrak{T}_T$, are a tight family in $C(I;M)$.

In the following sections, we will indeed show that all assumptions in Hypothesis 2 are satisfied in Stiefel manifolds.

Since this theorem is proved using Theorem 3, and this latter using Theorem 5, then all corollaries of Theorems 3 and 5 will hold also for Theorem 6.

We have, moreover, this result.

Corollary 9. *If the law of the starting r.v. $\mathfrak{X}_0 = \mathfrak{X}_0^\tau$ of the processes is invariant for right actions of $A \in O(p)$, then the law of $\mathfrak{X}^\tau$ is invariant for right actions of $A \in O(p)$; so, this is true for any limit point $\mathfrak{X}$ of $\mathfrak{X}^\tau$. So, the above can be interpreted as a result for Grassmann manifolds as well.*

7.3.1 | More on Geodesics

The results from [15] still hold, with minor adjustments in notation.

Lemma 8. *Given $x \in M = \mathbf{St}(p,V)$ and $v \in H$, we have that $v \in T_x M$ iff $x^\top v$ is an asymmetric matrix.*

Given $x \in M = \mathbf{St}(p,V)$ and $A \in O(p)$ orthogonal matrix, then $xA \in M$. The action $x \longmapsto xA$ is an isometry in H and hence in M.

Proposition 8 (Critical geodesics in (p, V)). *Let* $\mathbf{St}(p,V)$ *be endowed with the induced metric from* $H = V^p$. *Let* $\gamma\colon [0,1] \longrightarrow \mathbf{St}(p,V)$ *be a path. Then, the geodesic equation is* $\ddot\gamma + \gamma(\dot\gamma^\top \dot\gamma) = 0$. *Solutions to the geodesic equation exist for all time and are given by*

$$\left(\gamma(t)e^{At}, \dot\gamma(t)e^{At}\right) = (\gamma(0), \dot\gamma(0))\exp t\begin{pmatrix} A & -S \\ \mathbb{I} & A \end{pmatrix}, \tag{177}$$

where $\mathbb{I}$ *is the* $p \times p$ *identity matrix and* $A = \gamma(0)^\top \dot\gamma(0)$ *and* $S = \dot\gamma(0)^\top \dot\gamma(0)$.

Note that A is asymmetric and S is symmetric; moreover, $A = \gamma(t)^\top \dot\gamma(t)$ and $S = \dot\gamma(t)^\top \dot\gamma(t)$ are constant along the geodesic and $e^{At} \in O(p)$.

Further properties of infinite dimensional Stiefel and Grassmann manifolds are discussed in [16]. In particular, it is proven that any two points in those manifolds are connected by a minimal length geodesic.

We will now add more analysis to achieve the desired results.

Remark 13. Let $\lambda \in \mathbb{R}, \lambda \neq 0$, we fix $x \in M, v \in T_xM$ and set $(\gamma(0), \dot\gamma(0)) = (x, v)$. We define $A = x^\top v$ and $S = v^\top v$ as above. In the geodesic Equation (177), we multiply as follows:

$$\left(\gamma(t)e^{At}, \lambda\dot\gamma(t)e^{At}\right) = (x, \lambda v)\begin{pmatrix} \mathbb{I} & 0 \\ 0 & \lambda^{-1}\mathbb{I} \end{pmatrix}\exp\left(t\begin{pmatrix} A & -S \\ \mathbb{I} & A \end{pmatrix}\right)\begin{pmatrix} \mathbb{I} & 0 \\ 0 & \lambda\mathbb{I} \end{pmatrix}, \tag{178}$$

so

$$\begin{aligned}\left(\gamma(t)e^{At}, \lambda\dot\gamma(t)e^{At}\right) &= (x, \lambda v)\exp t\begin{pmatrix} A & -\lambda S \\ \lambda^{-1}\mathbb{I} & A \end{pmatrix} \\ &= (x, \lambda v)\exp t\lambda^{-1}\begin{pmatrix} \lambda A & -\lambda^2 S \\ \mathbb{I} & \lambda A \end{pmatrix}.\end{aligned} \tag{179}$$

We can use this relation as follows. Let now $\theta = \|v\|, \hat v = v/\theta$, $\hat A = x^\top \hat v$, and $\hat S = \hat v^\top \hat v$, then setting $\lambda = 1/\theta$,

$$\left(\gamma(t)e^{t\theta\hat A}, \theta^{-1}\dot\gamma(t)e^{t\theta\hat A}\right) = (x, \hat v)\exp t\theta\begin{pmatrix} \hat A & -\hat S \\ \mathbb{I} & \hat A \end{pmatrix}, \tag{180}$$

this formula decouples v into the initial direction $\hat v$ and the initial speed θ.

7.3.2 | Estimates

We define $\|v\|_K$ as in Definitions 11.

Lemma 9. *We recall the Definitions 13. For* $W \in \mathbb{R}^{k\times k}$, *we use the norm*

$$|W| = \sqrt{\sum_{i,j=1}^k |W_{i,j}|^2}, \tag{181}$$

and remark that

$$|WV| \le |V||W|; \tag{182}$$

we will use it with $k = p$ *or* $k = 2p$; *if* E *is a vector space with a scalar product, then for* $v, w \in E$, *we have*

$$|v^\top w| \le \|v\|_E \|w\|_E, \tag{183}$$

and for $v \in E^k$,

$$\|vW\|_{E^k} \le \|v\|_{E^k}|W| \tag{184}$$

by Cauchy–Schwarz inequality.

Lemma 10. *We consider the orthogonal projection* $\pi_T\colon H \longrightarrow H$ *to a hyperplane*

$$T = \{x \in H\colon \forall i \le v, \langle w_i, x\rangle_H = 0\}, \tag{185}$$

orthogonal to $w_1, \dots w_v \in H$, *where those vectors are orthogonal but not orthonormal; then,*

$$\pi_T v = v - \sum_{i=1}^v w_i \frac{\langle w_i, v\rangle_H}{\|w_i\|_H^2}. \tag{186}$$

we immediately note that if v, w_i *are in a vector subspace* E *of* H, *then* $\pi_T v$ *will be in the same* E. *Starting from Equation (186), we estimate*

$$\|\pi_T v\|_K \le \|v\|_K + \sum_{i=1}^v \|w_i\|_K \frac{\|v\|_H}{\|w_i\|_H}. \tag{187}$$

The tangent plane $T = T_xM$ *to the Stiefel manifold is such a plane, with*

$$w_i = (0, \dots 0, x_i, 0, \dots 0) \tag{188}$$

containing the ith column of x *in position ith; this for* $i = 1, \dots p$;

- *and then for* $i = p+1, \dots p(p+1)/2$,

$$w_i = (0, \dots 0, x_h, 0, \dots 0, -x_k, 0, \dots 0), \tag{189}$$

containing the hth column of x in position kth and vice versa and with a minus sign; so, by the diagonal structure (174) of K and by Equation (186) , we obtain the following: if $x \in M \cap K(H)$ *and* $v \in K(H)$, *then* $\pi_{T_xM} v \in K(H)$.

Moreover, such w_i *are mutually orthogonal, and* $\|w_i\|_H = 1$ *for* $i = 1, \dots p$, *while* $\|w_i\|_H = 2$, *otherwise, while* $\|w_i\|_K \le \|x\|_K$ *in all cases; so, in conclusion,*

$$\left\|\pi_{T_xM} v\right\| \le c_p\left(\|v\|_K + \|x\|_K\|v\|_H\right), \tag{190}$$

for a $c_p > 0$ *independent of* x, v. *This proves estimate (154) in Hypothesis 2.*

The above suggests that (154) in Hypothesis 2 may hold for other manifolds, as long as the embedding in H has finite codimension.

Lemma 11. *Let* $\tilde K\colon V \longrightarrow V$ *a linear continuous injective operator. We recall that* $H = V^p$ *and we defined* $K\colon H \longrightarrow H$ *in (174) as*

$$y = Kx \text{ when } y_i = \tilde K x_i. \tag{191}$$

So, K is a linear continuous operator and commutes with the right multiplication by matrixes

$$(Kx)A = K(xA). \tag{192}$$

There is a constant $c > 0$ *such that, for all* $x \in M \cap K(H), v \in T_xM \cap K(H)$ *and the geodesic with*

$$(\gamma(0), \dot{\gamma}(0)) = (x, v), \tag{193}$$

we have

$$\left(\|\gamma(1) - x - v\|_K + \|\dot{\gamma}(1) - v + Sx\|_K\right) \le c\left(\|x\|_K\|v\|_H^2 + \|v\|_K\|v\|_H\right)e^{c\|v\|_H}. \tag{194}$$

Note that c does not depend on K but only on p.

This proves (155) in Hypothesis 2.

Proof 7. We fix $x \in M, v \in T_xM$ and set $\theta = \|v\|, \hat{v} = v/\theta$. We will use the formula seen in Equation (180) with $t = 1$. We define

$$\hat{A} = x^\top \hat{v}, \hat{S} = \hat{v}^\top \hat{v},$$
$$Z = \begin{pmatrix} \hat{A} & -\hat{S} \\ \mathbb{I} & \hat{A} \end{pmatrix}, B = \begin{pmatrix} -\hat{A} & 0 \\ 0 & -\hat{A} \end{pmatrix}, \Theta = \begin{pmatrix} \mathbb{I} & 0 \\ 0 & \theta\mathbb{I} \end{pmatrix}, \tag{195}$$

so Formula (180) becomes

$$(\gamma(t), \dot{\gamma}(t)) = (x, \hat{v})\exp(t\theta Z)\exp(t\theta B)\Theta, \tag{196}$$

then

$$\begin{aligned}(\gamma(t), \dot{\gamma}(t)) &= (x, \hat{v})\left(\sum_{i\ge0, j\ge0} (t\theta)^{i+j}\frac{Z^i}{i!}\frac{B^j}{j!}\right)\Theta \\ &= (x, \hat{v})\left(\mathbb{I} + t\theta(Z + B) + \sum_{i,j,i+j\ge2} (t\theta)^{i+j}\frac{Z^i}{i!}\frac{B^j}{j!}\right)\Theta \\ &= \left(x + tv, v - t\theta^2 xS\right) + (x, \hat{v})\left(\sum_{i,j,i+j\ge2} (t\theta)^{i+j}\frac{Z^i}{i!}\frac{B^j}{j!}\right)\Theta.\end{aligned} \tag{197}$$

Now,

$$|\hat{A}| \le p, |\hat{S}| \le 1, |\hat{B}| \le 2p, |Z| \le \sqrt{3p + 1}, \tag{198}$$

and setting $t = 1$,

$$\begin{aligned}\sqrt{\|\gamma(1) - x - v\|_K^2 + \left\|\dot{\gamma}(1) - v + \theta^2 x\hat{S}\right\|_K^2} &\le \sqrt{\|x\|_K^2 + \|\hat{v}\|_K^2}\left|\sum_{i,j,i+j\ge2} \theta^{i+j}\frac{Z^i}{i!}\frac{B^j}{j!}\right||\Theta| \\ &\le \sqrt{\|x\|_K^2 + \|\hat{v}\|_K^2}\sum_{i,j,i+j\ge2} |\theta|^{i+j}\frac{|Z|^i}{i!}\frac{|B|^j}{j!}\sqrt{(1 + p\theta^2)} \\ &\le \sqrt{\|x\|_K^2 + \|\hat{v}\|_K^2}\theta^2 g(\theta).\end{aligned} \tag{199}$$

where we are using the fact (175) that left multiplication by K and right multiplication by a matrix are associative, and

$$\begin{aligned}g(s) &= \sum_{i,j,i+j\ge2} s^{i+j-2}\frac{(3p + 1)^{i/2}}{i!}\frac{(2p)^j}{j!} \\ &= \sum_{2\le n}\sum_{0\le k\le n}\frac{s^{n-2}}{n!}\left(\binom{n}{k}(3p + 1)^{k/2}\binom{n}{n-k}(2p)^{n-k}\right) = \sum_{2\le n}\frac{s^{n-2}}{n!}\left(\sqrt{3p + 1} + 2p\right)^n \\ &= \frac{1}{s^2}(\exp a - 1 - a) \text{ with } a = s\left(\sqrt{3p + 1} + 2p\right).\end{aligned} \tag{200}$$

□

The first and second assumptions in hypotheses in 2 are obviously true for Stiefel manifolds; indeed, the curvatures and second fundamental form are uniformly bounded since Stiefel manifolds are homogeneous spaces. Nonetheless, we can provide this estimate that satisfies (156).

Corollary 10. *There is a constant $c > 0$ such that for all $x \in M, v \in T_xM$ and the geodesic with*

$$(\gamma(0), \dot{\gamma}(0)) = (x, v), \tag{201}$$

we have

$$\|\gamma(1) - x - v\|_H \le c\min\left\{\|v\|_H, \|v\|_H^2\right\}. \tag{202}$$

Proof of Corollary 10. We note that the Stiefel manifold $\mathbf{St}(p, V)$ has diameter d so that for $v \ge d$, we can estimate

$$\|\gamma(1) - x - v\|_H \le d + \|v\|_H, \tag{203}$$

while for $v \le d$, we use the above Lemma 11 with K being the identity, recalling that $\|x\|_K = \sqrt{p}$ in this case. □

8 | Conclusions and Future Developments

We now know that, under appropriate hypotheses, the processes $\mathfrak{X}^\tau$ have narrow limit points $\mathfrak{X}$ when the partition τ becomes finer and finer; these $\mathfrak{X}$ are random functions in $C(\mathbb{R}^+; S)$ with $S = H$ or $S = M$ an embedded manifold.

There are multiple questions left unanswered, material for future research.

- Do the limit points enjoy some standard property? It seems plausible that they may enjoy some kind of Markov property, for example.
- Under which additional hypotheses can we say that there is a unique limit point?
- Can we then characterize the limit points as solutions to a kind of SDE?

These two questions are in synergy.

- Can we expand the results to more general cases of random walks, for example, where the constants in the hypothesis are not "uniform" but rather they may grow (e.g., be bounded by a function of the distance from a given point)?
- Consequently, are there other infinite dimensional manifolds where the present results hold true?

All the above questions are starting point for future research.

Acknowledgments

The author would like to thank Prof. Thomas G. Kurtz for the help in finding some references and Prof. Luciano Mari for helpful discussions.

This project was supported by the RICBASE_2022_MENNUCCI funding and by PRIN 2022 project 202244A7YL_002, Gradient Flows and Non-Smooth Geometric Structures with Applications to Optimization and Machine Learning.

No artificial intelligence tool was used in preparing this paper.

Funding

This project was supported by the RICBASE_2022_MENNUCCI funding and by PRIN 2022 project 202244A7YL_002, "Gradient Flows and Non-Smooth Geometric Structures with Applications to Optimization and Machine Learning". The funding RICBASE_2022_MENNUCCI refers to an Individual Research Grant allocated annually to Professors and Researchers by the Italian Ministry of University and Research through the National Fund for University Financing (FFO), pursuant to Law 537/1993. These funds are noncompetitive and are automatically assigned to faculty members to support their institutional research activities and basic scientific operations.

"PRIN" (Progetti di Ricerca di Rilevante Interesse Nazionale) are Italian government-funded grants for research projects, managed by the Ministry of University and Research (MUR). These funds are competitive.

Open access publishing facilitated by Scuola Normale Superiore, as part of the Wiley–CRUI-CARE agreement.

Disclosure

An earlier draft of this work was posted on the preprint server CVGMT at the author's webpage (https://cvgmt.sns.it/person/109/). Because the present version is a thorough revision, that preprint has been removed.

Conflicts of Interest

The author declares no conflicts of interest.

Data Availability Statement

Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

Endnotes

[1]Properly defined in Section 3.5.1.

[2][15] credits a personal communication by R. Lippert for the final closed form Formula (7.9).

[3]Other text calls this convergence in distribution or weak convergence.

[4]The constants c_3, c_t will appear again in subsequent hypotheses.

[5]The last two requests are loosely connected to what is explained in Example 3.8.13 in [13].

[6]See the version of Prokhorov's theorem in [9].

[7]See Definition 6.

[8]See Theorem 3.3 in [21] or Theorem 6.35 in [9].

[9]A possible proof may be found in Theorem 9.1 in [18] (this uses the linear approximations (71)). Note that *Wiener process* is also known as *Brownian motion* in some texts, as [10] or Chapter 12 in [22].

[10]Recall from Proposition 3 that $\mathcal{F}_t$ is the σ-algebra generated by $\mathfrak{X}_0$ and by Y_s for $s < t$.

[11]In the version in Theorem 3.8.2 in [13], or Theorem 3.5 in [23]; see Theorem 6.11 in [9] for convenience.

References

1. A. Mennucci, "Metrics of Curves in Shape Optimization and Analysis," in *Osher and Burger [2], Page 114. Lectures Given at the C.I.M.E. Summer School Held in Cetraro* (September 2008), 8–13.

2. S. Osher and M. Burger, eds., "Level Set and PDE Based Reconstruction Methods: Applications to Inverse Problems and Image Processing," in *Lecture Notes in Mathematics* (Berlin: Springer-Verlag, 2013), https://doi.org/10.1007/978-3-319-01712-9.

3. A. Srivastava, E. Klassen, S. H. Joshi, and I. H. Jermyn, "Shape Analysis of Elastic Curves in Euclidean Spaces," *IEEE Transactions on Pattern Analysis and Machine Intelligence* 33, no. 7 (July 2011): 1415--1428, https://doi.org/10.1109/TPAMI.2010.184.

4. L. Younes, P. W. Michor, J. Shah, and D. Mumford, "A Metric on Shape Space With Explicit Geodesics," *Atti Della Accademia Nazionale dei Lincei* 19, no. 1 (2008): 25–57, https://doi.org/10.4171/RLM/506.

5. L. Younes, "Computable Elastic Distances Between Shapes," *SIAM Journal on Applied Mathematics* 58, no. 2 (1998): 565–586, https://doi.org/10.1137/S0036139995287685.

6. G. Sundaramoorthi, A. Mennucci, S. Soatto, and A. Yezzi, "A New Geometric Metric in the Space of Curves, and Applications to Tracking Deforming Objects by Prediction and Filtering," *SIAM Journal on Imaging Sciences* 4, no. 1 (2011): 109–145, https://doi.org/10.1137/090781139.

7. S. Deiala, "Una Metrica Riemanniana Sullo Spazio Delle Curve e Applicazioni," *Master's thesis, Università di Pisa* (2010): https://etd.adm.unipi.it/theses/available/etd-10122010-204743/.

8. E. Bardelli and A. C. G. Mennucci, "Probability Measures on Infinite Dimensional Stiefel Manifolds," *Journal of Geometric Mechanics* 9, no. 3 (2017): 291–316, https://cvgmt.sns.it/paper/2844/, https://doi.org/10.3934/jgm.2017012.

9. A. C. G. Mennucci, "Some Fundamental Results on Probabilities and Stochastic Processes in Infinite Dimensional Spaces and Manifolds," (2024), https://cvgmt.sns.it/paper/6945/.

10. E. P. Hsu, "Stochastic Analysis on Manifolds," *Graduate Studies in Mathematics* (American Mathematical Society, 2002), https://books.google.it/books?id=GDEPCgAAQBAJ.

11. A. Grigor'yan, "Analytic and Geometric Background of Recurrence and Non-Explosion of the Brownian Motion on Riemannian Manifolds," *Bulletin of the American Mathematical Society* 36, no. 2 (1999): 135–249, https://doi.org/10.1090/s0273-0979-99-00776-4.

12. R. Azencott, "Behavior of Diffusion Semi-Groups at Infinity," *Bulletin de la Société Mathématique de France* 102 (1974): 193–240, https://doi.org/10.24033/bsmf.1778.

13. V. I. Bogachev, *Gaussian Measures, Volume 62 of Mathematical Surveys and Monographs* (American Mathematical Society, 1998).

14. C. Bessaga, "Every Infinite-Dimensional Hilbert Space Is Diffeomorphic With Its Unit Sphere," *Bulletin of the Polish Academy of Sciences. Mathematics* 14 (1966): 27–31.

15. A. Edelman, T. Arias, and S. Smith, "The Geometry of Algorithms With Orthogonality Constraints," *SIAM Journal on Matrix Analysis and Applications* 20, no. 2 (1998): 303–353, https://doi.org/10.1137/S0895479895290954.

16. P. Harms and A. Mennucci, "Geodesics in Infinite Dimensional Stiefel and Grassmann Manifolds," *Comptes Rendus Mathematique* 350, no. 15-16 (2012): 773–776, https://cvgmt.sns.it/papers/harmen10/, https://doi.org/10.1016/j.crma.2012.08.010.

17. G. Da Prato, *An Introduction to Infinite-Dimensional Analysis* (Springer, 2006).

18. P. Billingsley, "Convergence of Probability Measures," *Wiley Series in Probability and Statistics* (Wiley, 2013), https://books.google.it/books?id=6ItqtwaWZZQC.

19. N. Etemadi, "On Some Classical Results in Probability Theory," *Sankhya-Series A* 47, no. 2 (1985): 215–221, https://www.jstor.org/stable/25050536.

20. P. Billingsley, "Probability and Measure," *Wiley Series in Probability and Statistics* (Wiley, 1995), https://books.google.it/books?id=z39jQgAACAAJ.

21. G. Da Prato and J. Zabczyk, "Stochastic Equations in Infinite Dimensions," *Encyclopedia of Mathematics and Its Applications* (Cambridge University Press, 2014), https://books.google.it/books?id=bxkmAwAAQBAJ.

22. L. Breiman, *Probability* (Addison-Wesley, 1968).

23. N. Vakhania, V. Tarieladze, and S. Chobanyan, "Probability Distributions on Banach Spaces," *Mathematics and Its Applications* (Netherlands: Springer, 2012), https://books.google.it/books?id=JtLyCAAAQBAJ.

Appendix A. Useful Lemmas

In this section, we have collected the technical lemmas used here and there in the paper.

Lemma 12. *For E space with scalar product and $v, w \in E$,*

$$|v + w|^2 \le 2|v|^2 + 2|w|^2. \tag{A.1}$$

Proof 8.

$$|v + w|^2 \le (|v| + |w|)^2 \le 2|v|^2 + 2|w|^2. \tag{A.2}$$

□

Proof 9. In the following, monotonic means monotonically weakly increasing, that is, $s \le t \Longrightarrow g(s) \le g(t)$. □

Lemma 13. *Let $\beta \in \mathbb{R}$. Let*

$$t_0 = 0 < t_1 < t_2 < , \ \dots , \tag{A.3}$$

and let $\delta_n = t_{n+1} - t_n$. Suppose b_n is a real valued sequence with $b_n \ge \beta$ for all n. Suppose that

$$\varphi = \varphi(t, x) \colon [0, \infty) \times [\beta, \infty) \longrightarrow [0, \infty) \tag{A.4}$$

is a continuous nonnegative function such that $\varphi(\cdot, x)$ and $\varphi(t, \cdot)$ are monotonically increasing. Let $f \colon [0, \infty) \longrightarrow [\beta, \infty]$ be a solution of

$$\begin{cases} f'(t) = \varphi(t, f(t)), \\ f(0) = b_0, \end{cases} \tag{A.5}$$

possibly $f(t) = \infty$ for large t. If

$$b_{n+1} \le b_n + \varphi(t_n, b_n)\delta_n \tag{A.6}$$

holds, then

$$b_n \le f(t_n). \tag{A.7}$$

Proof 10. Proof by induction: note that f is monotonic since $f' \ge 0$ but then it is convex since f' is monotonic.

$$\begin{aligned} b_{n+1} &\le b_n + \varphi(t_n, b_n)\delta_n \le f(t_n) + \varphi(t_n, f(t_n))\delta_n \\ &= f(t_n) + f'(t_n)\delta_n \le f(t_{n+1}). \end{aligned} \tag{A.8}$$

□

Note that indeed (A.6) can be rewritten as

$$\frac{b_{n+1} - b_n}{t_{n+1} - t_n} \le \varphi(t_n, b_n). \tag{A.9}$$

Lemma 14. *Let*

$$t_0 = 0 < t_1 < t_2 < , \ \dots , \tag{A.10}$$

and let $\delta_n = t_{n+1} - t_n$; fix $c_7 > 0, c_8 > 0$. Suppose b_n is a real valued sequence with $b_0 = 0, b_n \ge 0$ that satisfies

$$b_{n+1} \le b_n + 2\delta_n\left(c_7\sqrt{b_n}\sqrt{t_n} + c_8 t_n\right); \tag{A.11}$$

then,

$$b_n \le (c_7 + 2c_8)g(t_n)\,\text{with}\,g(t) = \frac{e^{c_7 t} - 1 - c_7 t}{c_7^2}. \tag{A.12}$$

Moreover, we set

$$\hat{\varepsilon} = \frac{2}{c_7 + 2c_8}, \tag{A.13}$$

then

$$\forall n, t_n \le \hat{\varepsilon} \Longrightarrow b_n \le t_n^2(c_7 + c_8) \tag{A.14}$$

and note that $b_0 = b_1 = 0$.

Proof 11. We consider a solution of the ODE

$$\begin{cases} f'(t) = 2c_7\sqrt{f(t)}\sqrt{t} + 2c_8 t, \\ f(0) = 0, \end{cases} \tag{A.15}$$

that is, (A.5) for this special case. Since

$$\sqrt{ab} \le \frac{a + b}{2}, \tag{A.16}$$

then

$$f'(t) \le c_7 f(t) + (c_7 + 2c_8)t, \tag{A.17}$$

and substituting $f(t) = g(t)e^{c_7 t}$ and with some calculations, we obtain

$$f(t) \le (c_7 + 2c_8)g(t)\,\text{with}\,g(t) = \frac{e^{c_7 t} - 1 - c_7 t}{c_7^2}. \tag{A.18}$$

Since

$$f'(t) \ge 2c_8 t, \tag{A.19}$$

then

$$f(t) \ge c_8 t^2, \tag{A.20}$$

in particular for any solution, we have $f(t) > 0$ and $f'(t) > 0$ for $t > 0$. We have $f'(0) = 0$, so $f(t) \leq t$ for $t \leq \varepsilon$ with ε small; more precisely, note that $g(t)$ is convex and increasing and $g(0) = g'(0) = 0$, so we set ε to be the unique positive solution of

$$(c_7 + 2c_8)g(t) = t; \tag{A.21}$$

moreover, $g(t) \geq t^2/2$, so we know that

$$\varepsilon \geq \frac{2}{c_7 + 2c_8}. \tag{A.22}$$

Now, we set

$$\hat{s} = \sup\{s \geq 0 \colon t \leq s \Longrightarrow f(t) \leq t\}, \tag{A.23}$$

note that $\varepsilon \leq \hat{s} \leq 1/c_8$; for $t \in [0, \hat{s}]$,

$$f'(t) \leq 2(c_7 + c_8)t, \tag{A.24}$$

so

$$f(t) \leq (c_7 + c_8)t^2. \tag{A.25}$$

□

Lemma 15. *Let*

$$t_0 = 0 < t_1 < t_2 < , \ \dots , \tag{A.26}$$

and let $\delta_n = t_{n+1} - t_n$. Suppose b_n is a real valued sequence. If, for $c_7 > 0, c_8 \geq 0, c_9 \geq 0$,

$$b_{n+1} \leq b_n(1 + c_7\delta_n) + \delta_n(c_8 + c_9 t_n) \tag{A.27}$$

holds, then

$$\begin{aligned} b_n &\leq b_0 e^{c_7 t_n} + \left(e^{c_7 t_n} - 1\right)\left(\frac{c_8}{c_7} + \frac{c_9}{c_7^2}\right) - \frac{c_9 t_n}{c_7} \\ &= e^{c_7 t_n}\left(b_0 + \frac{c_8}{c_7} + \frac{c_9}{c_7^2}\right) - \left(\frac{(c_8 + c_9 t_n)}{c_7} + \frac{c_9}{c_7^2}\right). \end{aligned} \tag{A.28}$$

Proof 12. Indeed (A.27) can be rewritten as

$$\frac{b_{n+1} - b_n}{t_{n+1} - t_n} \leq c_7 b_n + c_8 + c_9 t_n, \tag{A.29}$$

and the associated differential equation is

$$f'(t) = c_7 f(t) + c_8 + c_9 t, \tag{A.30}$$

that has the solution

$$f(t) = e^{c_7 t} f(0) + \left(e^{c_7 t} - 1\right)\left(\frac{c_8}{c_7} + \frac{c_9}{c_7^2}\right) - \frac{c_9 t}{c_7}, \tag{A.31}$$

so this proves the result. □

Appendix B. Proofs

Proof of Proposition 1: If $c = 0$, then

$$\mathbb{E}[g(\delta\|Y\|)] = \delta^\alpha \mathbb{E}[\|Y\|^\alpha], \tag{B.1}$$

so we set

$$\tilde{c} = \mathbb{E}[\|Y\|^\alpha]. \tag{B.2}$$

Otherwise, we set

$$\tilde{c} = \mathbb{E}\left[\|Y\|^\alpha e^{cc_t\|Y\|}\right], \tag{B.3}$$

so

$$\mathbb{E}[g(\delta\|Y\|)] = \delta^\alpha \mathbb{E}\left[\|Y\|^\alpha e^{c\delta\|Y\|}\right] \leq \delta^\alpha \mathbb{E}\left[\|Y\|^\alpha e^{cc_t\|Y\|}\right] = \tilde{c}\delta^\alpha. \tag{B.4}$$

We recall Proposition 1.13 from [17].

Proposition 9. *Let $Y = N(a, Q)$ and*

$$\lambda_1 = \max_{\|x\| \leq 1} \langle x, Qx \rangle_H \tag{B.5}$$

be the highest eigenvalue of Q. Then, for $0 < \varepsilon < 1/\lambda_1$,

$$\int_H e^{\varepsilon\|x\|^2/2} \mathrm{d}\gamma(x) = \frac{\exp\left(-(1/2)\langle a, (1 - \varepsilon Q)^{-1} a \rangle_H\right)}{\sqrt{\det(1 - \varepsilon Q)}}, \tag{B.6}$$

whereas for $\varepsilon \geq 1/\lambda_1$, the integral is infinite.

Proof of Proposition 2: set $c_t = 1$ for simplicity. By the previous proposition, for any $\lambda > 0$,

$$\mathbb{E}\left[e^{\lambda\|Y\|_H}\right] < \infty. \tag{B.7}$$

For $k \in \mathbb{N}$ and $a > 0$, we have $a^k s^k \leq k! e^{as}$, hence choosing $k = \lceil \alpha \rceil$,

$$s^\alpha e^{sc} \leq \frac{k!}{a^k} e^{s(a+c)}, \tag{B.8}$$

so again, we define

$$\tilde{c} = \mathbb{E}[g(\|Y\|)] < \infty, \tag{B.9}$$

and we proceed as in the above proof of Proposition 1.

Proof of Proposition 4: By contradiction, suppose there is up to substituting $f(n)$ with $\cup_{j=0}^n f(j)$, we can suppose that f is monotonic. Let $c_t = 1$ for simplicity. We build iteratively $\tau \in \mathfrak{T}$ such that $\forall n, \tau \nsubseteq f(n)$, in this way. We will build a (nondecreasing) sequence $n_m \in \mathbb{N}$ such that $n_m \longrightarrow_m \infty$, and a sequence $t_0 = 0 < t_1 < \dots \in \tau$ satisfying the requisites in Definition 8. Let $t_0 = 0$, $t_1 = 1$ and $n_0 = n_1 = 0$; for $m \geq 1$ having chosen $t_m \in \tau$ and n_m, we look for $k > n_m$ such that there is a $t \in f(k) \setminus f(n_m) \wedge t \geq t_m + 1/2$.

- If there is no such k, we stop the iterative process by adding to τ an arbitrary sequence $t_{m+1} < t_{m+2} < \dots$ with $t_{m+j} \notin \cup_k f(k)$ and $1/2 < t_{m+j+1} - t_{m+j} < 1$; we set $n_{m+j} = n_m + j$, all that for $j \geq 0$.
- If there is such k, t, we add to τ an arbitrary sequence

$$t_{m+1} < t_{m+2} < , \ \dots , \ < t_{m+l} = t \tag{B.10}$$

such that

$$\frac{1}{2} < t_{m+j+1} - t_{m+j} < 1, \quad \text{for} j = 0, \ \dots \ l - 1; \tag{B.11}$$

then we set $n_{m+1} = \dots = n_{m+l} = k$; then, we repeat the iteration using $m + l$ as the new m.

In any case, we obtain that for infinitely many m, there is a l such that $t_{m+l} \notin f(n_m)$.